

Solving non-routine problems in medicine

Citation for published version (APA):

Rikers, R. M. J. P. (1999). *Solving non-routine problems in medicine*. [Doctoral Thesis, Maastricht University]. Universiteit Maastricht. <https://doi.org/10.26481/dis.19991124rr>

Document status and date:

Published: 01/01/1999

DOI:

[10.26481/dis.19991124rr](https://doi.org/10.26481/dis.19991124rr)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Solving Non-routine Problems in Medicine

Solving Non-routine Problems in Medicine

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan
de Universiteit Maastricht,
op gezag van de Rector Magnificus,
Prof. dr. A.C. Nieuwenhuijzen Kruseman,
volgens het besluit van het College van Decanen,
in het openbaar te verdedigen
op woensdag 24 november 1999 om 16.00 uur

door

Remigius Maria Johannes Petrus Rikers

geboren te Maastricht in 1967

Promotor

Prof.dr. H.G. Schmidt

Co-promotor

Dr. H.P.A. Boshuizen

Beoordelingscommissie

Prof.dr. G.J. Kok (voorzitter)

Dr. P.J. Braspenning

Prof.dr. H.F.M. Crombag

Prof.dr. W.M.G. Jochems (Open Universiteit, OTEC)

Dr. A.J.J.A Scherpbier

Omslag: Dr. M.M. van den Hurk, Veghel Design Group.

Productie: Datawyse | Universitaire Pers Maastricht

ISBN 90-9013251-1

Copyright © R.M.J.P. Rikers, 1999

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior permission of the author.

Voor mijn vader en moeder

Contents

Chapter 1

Medical Problem Solving: An Overview.....1

Chapter 2

The Robustness of Medical Expertise: Clinical Case Representations
by Medical Experts and Subexperts (Part 1).....21

Chapter 3

On the Constraints of Encapsulated Knowledge: Clinical Case
Representations by Medical Experts and Subexperts (Part 2).....37

Chapter 4

Knowledge Encapsulation and the Intermediate Effect.....53

Chapter 5

Effects of Clinical Case Priming on the Activation of Encapsulated
Knowledge: Differences between medical experts and subexperts...71

Chapter 6

Discussion and Conclusions.....93

Samenvatting & Conclusies.....107

Dankwoord115

Curriculum Vitae117

Chapter 1

Medical Problem Solving: An Overview

On the most general level, the study of expertise seeks to understand and account for what distinguishes outstanding individuals in a domain from less outstanding individuals in that domain, as well as from people in general. (Ericsson & Smith, 1991).

General introduction

What does it take to become a virtuoso on the piano, an Olympic skating star, a player of the Dutch national soccer team, or a chess player that is serious threat for Kasparov's crown? In other words, what does it take to become an expert in a domain? Surely, most people would agree that it takes a good set of genes and thousands of hours of practice. However, the widespread belief that a person must be innately talented in order to achieve an exceptional level of performance is disputed by biographical evidence. Over the years substantial evidence has accumulated that a person must practice many hours a week for at least 10 years before he or she can righteously claim: "I am an expert." This so-called "10-year rule" of Simon and Chase (1973) is supported by data from a wide range of domains: Swimming (Kalinowski, 1985), music (Sosniak, 1985), mathematics (Gustin, 1985), medical diagnosis (Boshuizen, 1989; Hobus, 1994; Lesgold, 1984; Patel & Groen, 1991a), and, of course, chess (Chase & Simon, 1973). In the domain of chess, for instance, Chase and Simon (1973) have shown that nobody had attained the level of a grandmaster with less than a decade of intense practice. The "myth of giftedness" is especially prominent within music, where children of the age of 6 perfectly play Beethoven's sonatas - at least that is what their parents claim. However, also within the musical domain a minimum of 10 years is required of practice to play or compose music that is truly outstanding. Note that it took the "classical" example of all prodigies, Wolfgang Amadeus Mozart, until his twenties before he reached his eminence (Ericsson & Lehmann, 1996).

In their biographical approach on the development of expertise, Ericsson, Krampe, and Tesch-Römer (1993) found that expert violinists prac-

ticed 7 hours per week by the age of 12. Their practice steadily increased, as they grew older, reaching more than 30 hours per week by the age of 20. Calculating their practice from the age of 3 to 20, they spent more than 10,000 hours of cumulative practice. Interestingly, according to Ericsson and colleagues (1993) sheer practice is not enough, the individual should engage in what they called *deliberate practice*. That is, the most effective learning requires a well-defined task with an appropriate difficulty level for the individual, informative feedback, and opportunities for repetition and corrections of errors. This regimen of effortful activities is typically begun in childhood. Playful activities with the primary goal of enjoyment are, obviously, not considered as deliberate practice. The central claim of this theory is that the level of performance of an individual is directly related to the amount of deliberate practice. Furthermore, deliberate practice is expected to have an important impact on the individual's cognitive system: There is a change in the structure and processes comprising the cognitive system, enabling a highly skilled individual to bypass the normal limits of information processing.

One of the first attempts to describe the changes in structure and process in the cognitive system as a result of practice has been made by Chase and Simon (1973) in their studies on chess expertise. Chase and Simon assumed that superior performance in chess, but also in other domains, was the result of vast amounts of knowledge and pattern-based retrieval acquired over many years of experience. In particular, they argued that high levels of chess performance were the result of large numbers of patterns in long-term memory (LTM) that served as retrieval cues to access chess moves appropriate for the corresponding chess position. Estimates of the magnitude of these patterns, usually referred to as *chunks*, ranges between 100,000 and a couple of a million (Richman, Gobet, Staszewski, & Simon, 1996). Furthermore, Chase and Simon (1973) asserted that these chunks mediated the experts' rapid encoding and retention of chess positions. Labels corresponding to particular chunks were held in short-term memory (STM) and subsequently used to retrieve each chunk in LTM. Although the Chase-Simon theory has proven to be successful to explain experts' superior performance, it is challenged by two empirical facts. First, the number of chunks recalled by experts exceeds the estimates of STM capacity (Broadbent, 1975; Chase & Ericsson, 1982; Ericsson & Kintsch, 1995; Ericsson & Staszewski, 1989; Glanzer & Razel, 1974). Second, experts' performance is only minimally affected by STM inference tasks (Charness, 1976; Ericsson & Staszewski, 1989; Ericsson & Kintsch, 1995). These two findings dispute the proposal that experts retrieve chunks of information held in LTM by maintaining appropriate cues in STM.

A theoretical alternative that can better explain the experts' superior performance of storage and retrieval of large amounts of information than the Chase-Simon theory is the so-called *Skilled Memory Theory* of Chase and Ericsson (1981, 1982).¹ The skilled memory theory gives a more detailed description of the knowledge experts possesses. Essentially it states that experts have acquired LTM encoding and retrieval skills that provides them with an expanded working memory capacity for materials within their domain of expertise. The speed and accuracy experts can access information in their LTM is normally associated with STM encoding and retrieval, but the capacity to store and retrieve information exceeds the capacity of STM by far. Experts' superior recall for familiar materials is viewed as a natural by-product of the experts' use of LTM to maintain information in an easily accessible state for processing in complex tasks (Ericsson & Staszewski, 1989).

The experts' efficient use of their LTM is characterized in terms of three principles: *Meaningful encoding* principle, *retrieval structure* principle, and the *speed-up* principle (Chase & Ericsson, 1981, 1982; Ericsson, Chase, & Faloon, 1980; Ericsson & Staszewski, 1989). The meaningful encoding principle (or mnemonic encoding) states that experts use their prior knowledge in order to durably encode information needed to successfully perform a routine task. Information is encoded in an well-organized network of concepts. The second principle, the retrieval structure principle, states that experts have developed mechanisms (i.e., retrieval structures) to facilitate the retrieval of information stored in LTM. The retrieval structures are used to encode information in LTM with cues that can later be regenerated to retrieve the stored information. That is, experts developed ways of organizing information they need to remember, and these organizational strategies do not depend on the specific items with which they are working (Payne & Wenger, 1998). For instance, in their study on the effects of practice on digit span, Chase & Ericsson (1981) showed that a participant who efficiently used retrieval structures increased his digit span by over 1000%. Finally, the speed-up principle states that as a result of practice encoding and retrieval in LTM will speed-up. Consequently, experts increase the amount of information they can access for processing far beyond the processing limitations associated with STM (Ericsson & Staszewski, 1989).

The principles of skilled memory theory are consistent with both earlier stage theories of skill acquisition (e.g., Fitts & Posner, 1967) and more contemporary theories on computational models (e.g., Anderson, 1983, 1987, 1990). Generally, these theories distinguish three stages in the development of

¹ Skilled memory theory has recently been extended into the theory of Long-term working memory (LTWM). This theory extends the basic ideas of *skilled memory theory* to a broader set of phenomena in memory in both text comprehension and expert performance (Ericsson & Kintsch, 1995).

skills. The first stage, the *cognitive stage*, is characterized by the processing of declarative knowledge or knowledge about basic facts about a task. Learners typically rehearse these facts as they perform their task. However, the knowledge acquired in the cognitive stage is quite inadequate for optimal performance. The second stage, tagged the *associative stage* of expertise development consists of eliminating errors in performance and strengthening the connections or associations among elements of task knowledge. During the associative stage so-called procedural knowledge becomes more and more important. The acquisition of procedural knowledge refers to *how* people perform a task. Finally, during the *autonomous stage*, which is essentially an extension of the associative stage, procedures become increasingly automated. For instance, Sweller, Mawer, and Ward (1983) studied the development of expertise in solving simple kinematics problems. They looked at how often subjects wrote down basic formulas involving velocity, distance, and acceleration (e.g., $V = A \times T$, where V is velocity, A is acceleration, and T is time). Initially, participants would write down all formulas so that they would not forget them. Gradually, however, as their competence increased, participants would only write down these equations with constants from the problem substituted for some of the variables (e.g., $V = 4 \times 18 \Rightarrow 72$). That is, in order to solve the problem they used the formula implicitly rather than to recall it explicitly.

In sum, on their "road to excellence" important cognitive changes occur as people develop their skills or gain expertise in a domain. However, there is still a large gulf between easily learned laboratory tasks, and skills exhibited by expert performers in the arts, sciences, and sports (Proctor & Dutta, 1995). This can be illustrated by the diagnostic skill in medicine, a domain that plays a central role in this thesis. The basic task in medical diagnosis is to place a patient into an appropriate disease category so that adequate treatment can be provided. However, the classification of a patient is a highly complex task and requires high levels of skilled performance. That is, the means available to determine the patient's condition are not clearly defined, but encompass all questions that could be asked and all laboratory tests that might be requested. Furthermore, the set of all possible diseases is vast and incomplete, and therefore, the selection of a treatment is often based on the best-fitting hypothesis than the only possible answer (Gilhooly, 1996). One might wonder how expertise development is possible at all in an "ill-structured" domain such as medicine. In order to shed some light on this issue, the next section will discuss the various phases a medical student has to go through in order to become a successful diagnostician.

The theory of Knowledge Encapsulation

What does it take to become an expert physician? Or, what are the changes a medical student has to go through in order to become an expert? To answer this question we have to start with a closer look at the program medical students have to go through in order to become a physician. During the first years in medical school (i.e., the pre-clinical years), students acquire large amounts of knowledge about the basic sciences. This kind of knowledge, which will further be referred to as biomedical knowledge, incorporates subjects such as physiology, anatomy, microbiology, and pathology. Essentially, biomedical knowledge refers to elements of causal mechanisms and characterizing patterns of perturbation in function and structure. The goal of teaching students first of all biomedical knowledge is to provide them with a scientific foundation for the tasks they have to face during clinical practice. Most traditional medical curricula are divided into pre-clinical years, which focus on biomedical knowledge, and a clinical, postgraduate program consisting of clinical courses and practicals. During these clinical years students are, often for the first time, confronted with real patients and they can witness in a real-life situation the impact of a disease on a patient's life. Clinical practice provides students with the opportunity to link their elaborate biomedical knowledge to their clinical experience.

The integration of biomedical knowledge with clinical knowledge is most clearly established by the theory of knowledge encapsulation (Schmidt & Boshuizen, 1992, 1993; Schmidt, Boshuizen, & Norman, 1992). Schmidt and Boshuizen assume that during the development of medical expertise, biomedical knowledge becomes *encapsulated* into clinically relevant concepts. That is, through extensive and repeated application of acquired knowledge (i.e., deliberate practice), and in particular through confrontation with real patients, their elaborate causal biomedical networks of disease are subsumed under high-level concepts with the same explanatory power.

According to Patel and colleagues this integration of clinical and biomedical knowledge is far from unproblematic. They assume that biomedical knowledge might even become *inert* in the course towards expertise, and it is, therefore, not explicitly used by expert physicians in their clinical case representation. Instead, expert physicians only use their clinical knowledge in order to build an adequate clinical case representation. In evaluating the results of their own studies, Patel and colleagues come to the conclusion that biomedical knowledge and clinical knowledge are "two different worlds," requiring different modes of reasoning and ways of structuring knowledge (Patel, Evans, & Groen, 1989). Clinical knowledge is based on a complex taxonomy, which relates signs and symptoms to the underlying pathology.

Biomedical knowledge, on the other hand, is based on general principles defining chains of causal mechanisms. Therefore, learning to explain how a set of signs and symptoms relate with a diagnosis may be very different from learning how to explain what causes a disease. Therefore, biomedical knowledge does not provide the axioms, the analogies, or the abstractions required to support clinical problem solving (Patel & Kaufman, 1995).

Although the integration of biomedical knowledge and clinical knowledge might be far from unproblematic, Schmidt and Boshuizen have shown that biomedical knowledge becomes, in the course towards expertise, gradually subsumed under clinically relevant encapsulating concepts.² Evidence to support their theory of knowledge encapsulation is found in the rather counter-intuitive phenomenon of *intermediate effect* in clinical case studies (e.g., Boshuizen, 1989; Claessen & Boshuizen, 1985; Patel & Medley-Mark, 1986; Schmidt & Boshuizen, 1992, 1993). In these studies participants of different levels of expertise were requested to study a text describing a patient's history, complaint(s), and some additional findings such as physical examination and laboratory data. After the text was removed, participants were asked to provide a diagnosis and to recall everything they remembered of the case. Finally, the participants were asked to explain the signs and symptoms displayed in the clinical case in terms of their disturbed processes (i.e., a pathophysiological explanation). The outcome of these studies was that participants of an intermediate level of expertise (e.g., advanced medical students) remembered more and gave more elaborate explanations of the clinical case descriptions than did expert physicians. So, rather than a monotonically increasing function of recall with increasing expertise, an inverted U-shaped relation was found. According to Schmidt and Boshuizen (1992, 1993), expert physicians, in contrast to novices and advanced students, introduce many encapsulating concepts in their recall and pathophysiological explanations, and thereby producing highly condensed protocols. For instance, Schmidt and Boshuizen (1992) asked participants of different levels of expertise to diagnose a clinical case description of *acute bacterial endocarditis* (see Table 1.1).

²Another position on the role biomedical knowledge is that it becomes *rudimentary*. This position implies that biomedical knowledge is acquired and applied in the early stages of the development of expertise, but will become gradually less important or even inaccessible. However, empirical evidence to support this position is not very strong and will therefore not be discussed any further (see for a more detailed discussion of the different positions, Boshuizen & Schmidt, 1992)

Table 1.1**Clinical case description of Acute Bacterial Endocarditis**

This 27-year-old unemployed male was admitted to the emergency room with the complaint of shaking chills and fever of four days' duration. He took his own temperature and it was recorded at 40° C. on the morning of his admission. The fever and chills were accompanied by sweating and a feeling of prostration. He also complained of some shortness of breath when he tried to climb the two flights of stairs in his apartment. Functional inquiry revealed a transient loss of vision in his right eye which lasted approximately 45 seconds on the day before his admission to the emergency ward.

Physical examination revealed a toxic looking young man who was having a rigor. His temperature was 41° C., pulse 120, BP 110/40. Mucus membranes were pink. Examination of his limbs showed puncture wounds in his left antecubital fossa. The patient volunteered that he had been bitten by a cat at a friend's house about a week before admission. There were no other skin findings. Examination of the cardiovascular system showed no jugular venous distention, pulse was 120 per minute, regular, equal and synchronous. The pulse was also noted to be collapsing. The apex beat was not displaced. Auscultation of his heart revealed a 2/6 early diastolic murmur in the aortic area and funduscopy revealed a flame shaped hemorrhage in the left eye. There was no splenomegaly. Urinalysis showed numerous red cells but there were no red cell casts. [Source: Patel, V.L., & Groen, G.J. (1986). Knowledge-based solution strategies in medical reasoning, *Cognitive Science*, 10, 91-116].

One expert physician recalled this case of acute bacterial endocarditis as follows: "The patient is a young man with a high fever who presents a *septic syndrome*. This suggests *drug use*. He shows signs of *thromboemboli*, due to an affected heart valve. The *tachycardia* fits with an associated *aorta vitium*." The expert's recall almost entirely consists of summaries of the information presented in the text. For instance, "septic syndrome" summarizes or encapsulates the following items within the text: High fever, shaking chills, sweating, feelings of prostration, shortness of breath, and high pulse rate. That is, these signs and symptoms activate a pre-existing encapsulating concept that is linked with these items. Advanced medical students, on the other hand, reproduced the information provided in the text almost literally, almost without summaries. Therefore, intermediates' free recall is not really better than that of expert physicians, unless one wants to equate more with better (Patel & Groen, 1991b).

The difference in clinical case processing by medical experts and students is perhaps most clearly demonstrated by their pathophysiological explanations of a clinical case. In a pathophysiological explanation participants are instructed to provide an explanation for the signs and symptoms displayed in a clinical case. For instance, Schmidt and Bosshuizen (1993)

instructed expert physicians and medical students to explain the signs and symptoms of the clinical case bacterial endocarditis described above. Table 1.2 contains an example of an actual verbatim protocol from a fourth-year medical student.

Table 1.2

Verbatim protocol fourth-year medical student

"Probably this young man is an intravenous drug addict and he is bitten by a cat. His resistance is not too good, so probably his immune system has not been able to eliminate the bacteria that have entered the man's body through the cat bite. As a consequence, his blood is invaded with bacteria. These bacteria produce toxins that are rendered harmless by antibodies. The complete system is activated and through this mechanism vasoactive substances are released, such as histamine and serotonin, etc. This is what is called the "hot phase" of the septic shock. The temperature control center is disturbed as well and is reset at a higher point. The body loses much heat and slowly the "cold phase" of the anaphylactic shock is entered. This is characterized by shivers, pallor and a coldness of the extremities of the body, disseminated intravascular coagulation (because of a vicious circle in which several influences play a role, such as deteriorated oxygen supply, toxins, endothelium damage caused by a hypoxemia etc.), I think, ablatio retinae can be caused by this as well. At the moment of entrance his blood pressure was low (namely 110/40 mmHg), then he was in the hot phase, in which reduced filling of the vascular bed caused this hypotension. Probably, the red cells in his urine can be explained by the disseminated intravascular coagulation that, accompanied with too much use of thrombocytes, causes a hemorrhagic diathesis; so this is caused by the consumption of platelets. In the end all organs are affected."

[Source: Schmidt, H.G., Boshuizen, H.P.A., & Norman, G.R. (1992). Reflections on the nature of expertise in medicine. In: *Deep models for medical knowledge engineering*. E. Keravnou. (Ed.). Elsevier Science Publishers].

The student used 42 propositions in order to explain the patient's signs and symptoms. The expert physician, on the other hand, only needed 10 propositions to do the same thing. So, instead of providing an elaborate explanation for each sign and symptom of the clinical case, as the student did, the expert only used a few highly inclusive or encapsulated concepts and, thereby, skipping the student's detailed steps (cf. Blessing & Anderson, 1996; Koedinger & Anderson, 1990). The difference in case processing might also be illustrated if the protocol of the expert physician and student are translated into a so-called semantic network representation. Figure 1.1 shows a semantic network representation that is derived from the student's explanation. Each proposition of the student's protocol is represented by two concepts that are linked to each other. The contrast with the expert physician's semantic network is striking (see Figure 1.2).

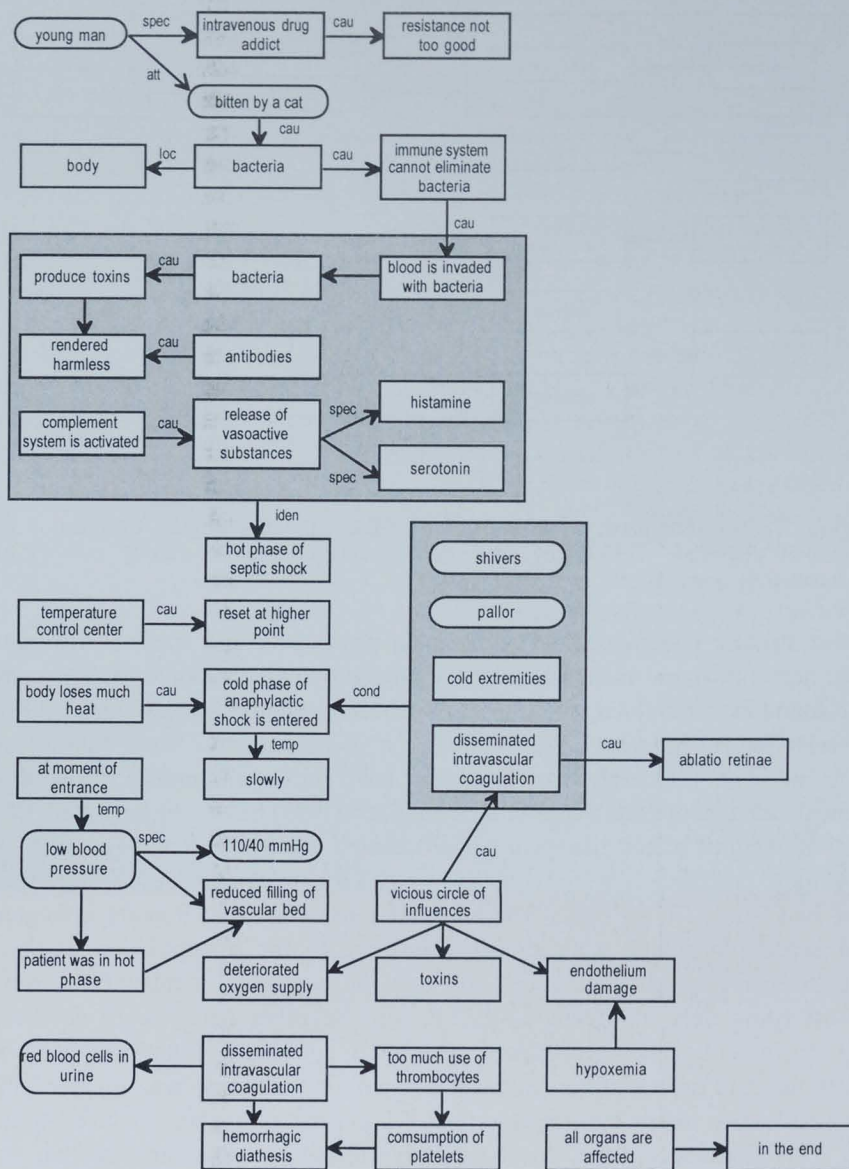


Figure 1.1. A semantic network representation of a fourth-year medical student's pathophysiological explanation. Frames with rounded corners represent signs and symptoms present in the case, and frames with rectangular corners represent explanatory concepts. The links between concepts are qualified as *cau.* (causal), *att.* (attribute), *loc.* (location), *spec.* (specification), *temp.* (temporal), *cond.* (conditional), *iden.* (identity).

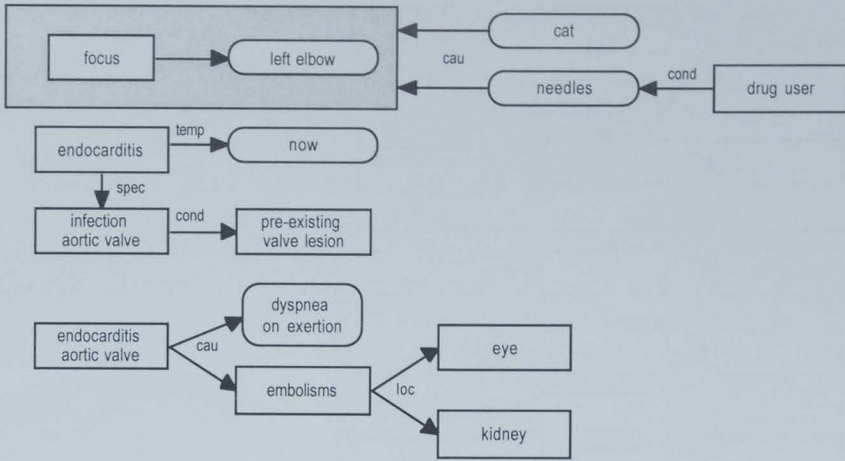


Figure 1.2. Semantic network representation of the pathophysiological explanation of an expert physician. This highly condensed explanation almost entirely consists of encapsulating concepts.

These striking differences between expert physicians and medical students were not only found in recall and pathophysiological explanations, but were also found in think-aloud studies (e.g., Boshuizen, Schmidt, & Coughlin, 1987; Schmidt & Boshuizen, 1993). However, it is important to note that these differences only emerge when students have sufficient time to process a clinical case elaborately. Schmidt & Boshuizen (1993) have demonstrated that the intermediate effect disappears if processing time is severely reduced. In this study, participants were required to recall a clinical case description after they had studied it for either 3.30 min, 1.15 min, or 30 sec. Results indicated that expert physicians were almost insensitive to the manipulation of time: The extensiveness of their recall and pathophysiological explanations were almost completely unaffected. Students on the other hand, were severely impaired if study time was reduced. The intermediate effect that was present under the 3.30 min condition in recall and pathophysiology was replaced by the well-known expert superiority effect if processing time was reduced to 30 sec. This finding confirms the hypothesis of encapsulation theory that by reducing processing time, students do not have sufficient time to engage in deep pathophysiological reasoning in order to understand the case and lose therewith their advantage in term of text recall. Experts, however, construct a presentation that only contains a small number of highly inclusive encapsulating concepts and are therefore relatively insensitive for manipulations of time. In other words, if students have sufficient time, they

process and hence report too much “garbage,” whereas experts do not (Patel & Groen, 1991b).

In sum, research using free recall, think-aloud protocols, and pathophysiological explanations produced clear results in regard to the role of encapsulated knowledge within the expert physicians' domain of expertise. Experts recall a clinical case mainly using concepts not found in the text itself, but summarizing most of the information from the text. Furthermore, their pathophysiological explanations also contain more summarizing or encapsulating concepts than those of medical students do. Therefore, although students remember more and give elaborate explanations than do expert physicians, this does not necessarily mean that their recall and pathophysiological explanations are better.

The nature of sub-expertise

In their book “The Nature of Expertise” Chi, Glaser, and Farr (1988), gave an overview of the characteristics of an expert in a domain based on findings of 40 years of research in this area. For instance, experts excel mainly in their domain of expertise; experts are faster than novices at performing skills within their domain; experts perform their task (almost) error free; experts have superior short-term and long-term memory; experts' problem representation is deeper, more principled, than that of novices, who tend to create superficial representations of a problem. From this list of expertise attributes one might conclude that we have a clear picture of the nature of expertise. However, it is important to note that most of these characteristics are derived from research with experts performing tasks that are highly routine. Less clear, however, is what happens to the experts' performance if experts are confronted with problems that are less routine or even non-routine. Do experts still show that same superior performance as before, or does their performance deteriorate dramatically?

In order to shed some light on this issue we have to go back to the roots of expertise research. It was in the 1940s that Adriaan de Groot, a Dutch psychologist and chess player, did his famous studies on chess players. In one of his studies chess players had to recall a position after seeing it for just a few seconds. What he found was that grandmasters were able to recall the position perfectly, and less “gifted” players were able to reproduce the position only partly (De Groot, 1946, 1965). In a subsequent experiment, Chase and Simon (1973) found that when the chess position was changed into a randomized position, the clear and significant differences between strong and weak players disappeared. This finding seems to suggest that expert chess

players evaluating normal chess positions detect common patterns of pieces, because if they would store individual pieces instead of patterns, randomization of the board would have no or only limited effects

In an attempt to adapt the methodology used by Chase and Simon (1973) to medicine, Coughlin and Patel (1987) used clinical case descriptions in which the information was either in the standard patient presentation format or it was randomly reordered. These cases were given to a group of expert physicians and medical students. Participants were instructed to read a case for a period of four minutes and were subsequently asked to recall the text they just read. The results of this study resemble those of Chase and Simon (1973) in chess. That is, experts perform significantly better than do medical students if they have to recall a case that was presented in a normal patient presentation format. However, when the clinical case is presented randomly, experts' performance deteriorates dramatically to the level of the medical students. Similarly, in the medical domain of nephrology, Norman, Brooks, and Allen (1989) presented protocols of laboratory data to medical experts (nephrologists), intermediates (second-year medical students), and research assistants who had no knowledge of relevant biochemistry. The protocols contained numerical values of laboratory data and were based on real patients or on randomized values. Participants were instructed to study the protocols and subsequently they had to recall the laboratory data as accurate as possible. The results showed almost the same pattern as Chase and Simon's chess studies: The number of laboratory data recalled increased with level of expertise if the data were based on a real patient. However, if random values were used, instead of realistic data, the expert superiority effect completely disappeared. Expert recall performance was as good or as bad as that of medical students or research assistants. As has been outlined before, findings in chess were explained in terms of the disruption of familiar patterns used by chess experts to recall a chess position. This explanation of disrupting familiar patterns might also be valid within the medical domain. This time, however, the patterns are the expert physicians' prototypical representation of familiar cases that are disturbed (e.g., Coughlin & Patel, 1987; Hobus, 1994).

It might be argued that, although the presentation of a clinical case in a random order is a non-routine task for expert physicians, it is also to a certain extent an artificial task. That is, it does not really answer another key question of how expert physicians deal with *realistic* problems that they did not encounter on a regular basis. For instance, what differences exist between expert physicians solving realistic problems within their specialty as compared to expert physicians solving realistic problems outside their specialty (in the latter case experts are dubbed *subexperts*)? As has been

outlined before, according to Schmidt and Boshuizen (1992, 1993), expert physicians confronted with problems within their specialty rely on their encapsulated knowledge in order to build an adequate clinical case representation. Expert physicians confronted with problems that are outside their specialty however, can not, or only to a limited extent, rely on encapsulated knowledge. In that case, expert physicians are expected to relapse into an elaborate biomedical-processing mode, which is typical of advanced medical students. Schmidt and Boshuizen have postulated that the different knowledge structures acquired in the course of the development towards expertise are stored in the expert's long-term memory, and are still accessible when more recently acquired, encapsulated structures, fail to produce an adequate representation of a clinical problem (Boshuizen & Schmidt, 1992; Schmidt & Boshuizen, 1992, 1993; Schmidt, Boshuizen, & Norman, 1992).

Evidence to support the relapse of expert physicians towards a biomedical-processing mode is provided by Patel, Arocha, and Groen (1986, 1990). They compared the performance of board certified medical specialists from two specialties (cardiology and endocrinology) in diagnosing clinical problems in the domains of cardiology and endocrinology. The specialists had to diagnose, recall, and explain clinical cases within and outside their domain of expertise. As might be expected, diagnostic accuracy decreased as a transition was made from the level of expert to that of subexpert. More interestingly, however, was that the specialists who had to explain a problem in their domain of expertise, primarily focused on patient-oriented components (i.e., clinical aspects) of the problem. Confronted with a clinical case outside their domain of expertise, their focus in their explanations shifted towards biomedical aspects of the problem. Surprisingly, Patel and colleagues (1986, 1990) found no significant differences between experts and subexpert in the number of propositions recalled. This finding is at variance with most recall studies across different domains who generally demonstrate the superior performance of experts on meaningful stimuli (e.g., Charness, 1979; Ericsson, 1985; Jensen, 1990; McKeithen, Reitman, Rueter, & Hirtle, 1981; Reitman, 1976; Spilich, Vesonder, Chiesi, & Voss, 1979). Patel and colleagues explained this finding by assuming that experts and subexperts alike have developed a kind of *generic expertise* for relevant information in a clinical case. Generic expertise is concerned with the acquisition of adequate case representations, implying the existence of some schema or macrostructure representation (Kintsch, 1988; Van Dijk & Kintsch, 1983). This representation is based on the physician's experience in practical problems, and guarantees the retention of the crucial facts of a clinical case within and outside the domain of expertise. Subexperts are therefore still able to recall the essential information of a clinical case. Expert physicians, on the other hand, have also acquired the so-

called *specific expertise*, which is essential to achieve a high diagnostic accuracy. They possess the necessary domain knowledge that enables them to link nonsalient cues or loose ends in the clinical case to the main diagnosis, and thereby producing a higher diagnostic accuracy (Patel & Groen, 1991a)

The present studies

Clinical case studies have repeatedly demonstrated the central role of encapsulated knowledge within the physician's domain of expertise (e.g., Boshuizen, 1989; Schmidt & Boshuizen, 1993; Schmidt, Boshuizen, & Norman, 1992). Furthermore, the intermediate effect, as it was found in clinical case studies comparing the performance of expert physicians with that of students, is probably the most striking exponent of the role of encapsulated knowledge. In order to disentangle the issue of the role of encapsulated knowledge outside the physicians' domain of expertise, the present thesis compared the performance of medical experts with that of subexperts.

Based on the theory of knowledge encapsulation it is predicted that encapsulated knowledge is mainly confined to the physician's domain of expertise. That is, a clinical case representation of experts within their specialty is characterized by encapsulating concepts. Subexperts (i.e., expert physicians diagnosing clinical cases outside their specialty), on the other hand, will relapse into an elaborate biomedical processing mode, which is characteristic for advanced medical students.

In Chapter 2 a study is presented that compares the performance of cardiologists and pulmonologists diagnosing cases within and outside their specialty. That is, cardiologists and pulmonologists fulfilled both roles of experts and subexperts in this study. It was hypothesized that if elaborate biomedical knowledge is encapsulated under more high-level concepts, experts would be more accurate and faster diagnosing cases within their specialty than outside. Furthermore, their recall and pathophysiological explanations should be more compendious within their specialty than outside, because they are highly encapsulated.

In Chapter 3 a study is reported that is a replication of the first study described in Chapter 2. However, there are four important differences with the first study. First, instead of pulmonologists, neurologists participated. Second, all cases were not as long and as complex as those described in Chapter 2, and were within the domain of cardiology. Third, all cases were considered to be routine problems for the cardiologists (i.e., the experts) and non-routine problems for the neurologists (i.e., subexperts). Finally, processing time was restricted. It was predicted that if encapsulated knowledge is

confined to the physician's domain of expertise, cardiologists' recall and pathophysiological explanations should be more compendious than those of neurologists, who have to resort to their biomedical knowledge in order to understand the signs and symptoms of the cases.

In Chapter 4 the "intermediate effect" is revisited, but in contrast to previous clinical case studies demonstrating the intermediate effect, cases were used outside the expert physician's specialty. Second-year, and sixth-year medical students, and neurologists had to diagnose clinical cases that were non-routine problems for the neurologists (i.e., subexperts). It was predicted that experts outside their specialty relapse into an elaborate processing mode and, therefore, the intermediate effect is expected to disappear. That is, subexperts' diagnostic accuracy, recall and pathophysiological explanations should not significantly differ from those of advanced medical students.

In Chapter 5 a study is reported that introduces a completely different approach within clinical case studies: The *priming paradigm*. Instead of the "traditional" measures of free recall and pathophysiological explanations, experts had to evaluate different types of items on their relatedness to the clinical case after studying a clinical case description. The priming approach makes it possible to probe the clinical case representation more directly. This chapter describes two experiments using the priming approach. Experiment 1 involved fourth-year medical students, subexperts, and expert physicians who had to study clinical cases within the domain of internal medicine. Subsequently, participants had to evaluate three types of items, displayed on a computer screen, as fast as possible: Literally stated signs and symptoms, inferred encapsulating items, and filler items that were completely unrelated to the clinical case. It was predicted that experts should be faster and more accurate than subexperts and medical students. Especially the expert physicians' response to encapsulating items should lead to speeded responses. Subexperts and medical students, on the other hand, are expected to show no significant differences as a result of the subexperts' relapse. Experiment 2 was to a large extent a replication of Experiment 1. However, the important distinction was that instead of signs and symptoms biomedical items were used. Fourth-year medical students, subexperts, and expert physicians had to evaluate these items as fast and as accurate as possible. Again, experts were expected to outperform both other groups. Furthermore, if biomedical knowledge is only of secondary importance within the expert physician clinical case representation, encapsulating items should especially lead to speeded responses.

In final chapter of this thesis, a summary is provided of all described studies, and implications for encapsulation theory will be discussed. Finally, suggestions for further research will be provided.

References

- Anderson, J.R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Anderson, J.R. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review*, 94, 192-210.
- Anderson, J.R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.
- Blessing, S. B., & Anderson, J. R. (1996). How do people learn to skip steps? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22 (3), 576-598.
- Boshuizen, H. P. A. (1989). *De ontwikkeling van medische expertise; een cognitief-psychologische benadering*. [On the development of medical expertise; a cognitive psychological approach]. Doctoral dissertation. Haarlem: Thesis.
- Boshuizen, H. P. A., & Schmidt, H.G. (1992). On the role of biomedical knowledge in clinical reasoning by experts, intermediates and novices. *Cognitive Science*, 16, 153-184.
- Boshuizen, H. P. A., Schmidt, H. G., & Coughlin, L. D. (1987). *On-line representation of a clinical case and the development of expertise*. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, DC.
- Broadbent, D.E. (1975). The magical number seven after fifteen years. In A. Kennedy & A. Wilkes (Eds.), *Studies in long-term memory*, New York: Wiley.
- Charness, N. (1976). Memory for chess positions: Resistance to interference. *Journal of experimental psychology: Human learning and memory*, 2, 641-653.
- Charness, N. (1979). Components of skill in bridge. *Canadian Journal of Psychology*, 33, 1-16.
- Chase, W.G., & Ericsson, K.A. (1981). Skilled memory. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 149-189). Hillsdale, NJ: Erlbaum.
- Chase, W.G., & Ericsson, K.A. (1982). Skill and working memory. In G.H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 16, pp. 1-58). San Diego, CA: Academic Press.
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55-81.
- Chi, M. T. H., Glaser, R. & Farr, M. (Eds.). (1988). *The nature of expertise*. Hillsdale, NJ: Erlbaum.
- Claessen, H. F. A., & Boshuizen, H. P. A. (1985). Recall of medical information by students and doctors. *Medical Education*, 19, 61-67.

- Coughlin, L. D., & Patel, V. L. (1987). Processing of critical information by students and doctors *Medical Education*, 62, 818-828.
- De Groot, A.D. (1946). *Het denken van den schaker* [Thinking processes in chess players]. Den Haag, The Netherlands: Noord-Holland.
- De Groot, A. D. (1965). *Thought and choice in chess*. The Hague, The Netherlands: Mouton.
- Ericsson, K. A. (1985). Memory skill. *Canadian Journal of Psychology*, 39(2), 188-231.
- Ericsson, K. A., Chase, W. G., & Faloon, S. (1980). Acquisition of a memory skill. *Science*, 208, 1181-1182.
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102(2), 211-245.
- Ericsson, K. A., Krampe, R. Th., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 3, 363-406.
- Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, 47, 273-305.
- Ericsson, K. A., & Smith, J. (Eds.). (1991). *Toward a general theory of expertise: prospects and limits*. Cambridge: Cambridge University Press.
- Ericsson, K.A., & Staszewski, J. (1989). Skilled memory and expertise: Mechanisms of exceptional performance. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp. 235-267). Hillsdale, NJ: Erlbaum.
- Fitts, P., & Posner, M.I. (1967). *Human performance*. Monterey, CA: Brooks/Cole.
- Gilhooly, K.J. (1996). *Thinking: Directed, undirected and creative* (3rd ed.). London, UK: Academic Press.
- Glanzer, M., & Razel, M. (1974). The size of the unit in short-term storage. *Journal of verbal learning and verbal behavior*, 13, 114-131.
- Gustin, W.C. (1985). The development of exceptional research mathematicians. In B.S. Bloom (Ed.). *Developing talent in young people* (270-231). New York: Ballantine Books.
- Hobus, P. P. M. (1994). Expertise van huisartsen. Praktijkervaring, kennis en diagnostische hypothesevorming [Family physician's expertise: Practical experience, knowledge and formation of diagnostic hypotheses]. Doctoral dissertation. Maastricht University Press, the Netherlands.
- Jensen, A.R. (1990). Speed of information processing in a calculating prodigy. *Intelligence*, 14, 259-274.
- Joseph, G. M., & Patel, V. L. (1990). Domain knowledge and hypothesis generation in diagnostic reasoning. *Journal of Medical Decision Making*, 10, 31-46.
- Kalinowski, A.G. (1985). The development of Olympic swimmers. In B.S. Bloom (Ed.). *Developing talent in young people* (139-192). New York: Ballantine Books.

- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95(2), 163-182.
- Koedinger, K. R., & Anderson, J. R. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science*, 14, 511-550.
- Lesgold, A. M. (1984). Acquiring expertise. In J.R. Anderson & S.M. Kosslyn (Eds.), *Tutorials in learning and memory: Essays in honor of Gordon Bower*, (pp 31-60). San Francisco, CA: Freeman.
- McKeithen, K. B., Reitman, J. S., Rueter, H. H., & Hirtle, S. C. (1981). Knowledge organization and skill differences in computer programmers. *Cognitive Psychology*, 13, 307-325.
- Norman, G. R., Brooks, L. R., & Allen, S. W. (1989). Recall by expert medical practitioners and novices as a record of processing attention. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 15, 1166-1174.
- Patel, V. L., Arocha, J. F., & Groen, G. J. (1986). Strategy selection and degree of expertise in medical reasoning. In C. Clifton (Ed.), *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*, (pp. 780-791). Hillsdale, NJ: Erlbaum.
- Patel, V. L., Evans, D. A., & Groen, G. J. (1989). Biomedical knowledge and clinical reasoning. In D. A. Evans & V. L. Patel (Eds.), *Cognitive science in medicine: Biomedical modelling*, (pp. 53-112). Cambridge, MA: MIT Press.
- Patel, V. L., & Groen, G. J. (1991a). The general and specific nature of medical expertise: A critical look. In A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits*, (pp. 93-125). New York: Cambridge University Press.
- Patel, V. L., & Groen, G. J. (1991b). Developmental accounts of the transition from medical student to doctor: Some problems and suggestions. *Medical Education*, 25(6), 527-535.
- Patel, V. L., Groen, G. J., & Arocha, J. F. (1990). Medical expertise as a function of task difficulty. *Memory and Cognition*, 18(4), 394-406.
- Patel, V. L., & Kaufman, D. R. (1995). Clinical reasoning and biomedical knowledge: Implications for teaching. In J. Higgs & M. Jones (Eds.), *Clinical reasoning in the health professions*, (pp. 117-128). Oxford, UK: Butterworth Heinemann.
- Patel, V.L., & Medley-Mark, V. (1986). Relationship between representation of textual information and underlying problem representation in medicine. *Cognitive research series*. Montreal: McGill University, Centre for Medical Education. (CME86-CS1).
- Payne, D. G., & Wenger, M. J. (1998). *Cognitive Psychology*. New York: Houghton Mifflin Company.
- Proctor, R.W., & Dutta, A. (1995). *Skill acquisition and human performance*. Thousand Oaks, CA: Sage Publications.
- Reitman, J. S. (1976). Skilled perception in Go: Deducing memory structures from inter-response times. *Cognitive Psychology*, 8, 336-356.

- Richman, H.B., Gobet, F., Staszewski, J.J., & Simon, H.A. (1996). Perceptual and memory processes in the acquisition of expert performance: The EPAM model. In K.A. Ericsson (Ed), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Schmidt, H. G., & Boshuizen, H. P. A. (1992). Encapsulation of biomedical knowledge. In D. A. Evans & V. L. Patel (Eds.), *Advanced models of cognition for medical training and practice*. New York, NY: Springer Verlag.
- Schmidt, H. G., & Boshuizen, H. P. A. (1993). On the origin of intermediate effects in clinical case recall. *Memory and Cognition*, 21(3), 338-351.
- Schmidt, H. G., Boshuizen, H. P. A., & Norman, G. R. (1992). Reflections on the nature of expertise in medicine. In E. Keravnou (Ed.), *Deep models for medical knowledge engineering*. Amsterdam, The Netherlands: Elsevier.
- Sosniak, L.A. (1985). Learning to be a concert pianist. In B.S. Bloom (Ed.). *Developing talent in young people* (19-67). New York: Ballantine Books.
- Spilich, G.J., Vesonder, G.T., Chiesi, H.L., & Voss, J.F. (1979). Text processing of domain-related information for individuals with high and low domain knowledge. *Journal of Verbal Learning & Verbal Behavior*, 18, 275-290.
- Sweller, J., Mawer, R.F., & Ward, M.R. (1983). *Development of expertise in mathematical problem solving*. *Journal of Experimental Psychology: General*, 112, 639-661.
- Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Chapter 2

The Robustness of Medical Expertise: Clinical Case Representations by Medical Experts and Subexperts (Part 1)¹

Medical specialists, confronted with problems within their domain of expertise, do not rely on intentional causal reasoning, using explicitly principles or rules of their domain. Rather, reasoning has become an automatic process, using knowledge in an encapsulated mode. Less clear, however, is what happens when medical specialists encounter problems outside their specialty. The purpose of the present study is to shed light on this issue. The study involved cardiologists and pulmonologists who had to evaluate four clinical cases. Two cases were within the domain of cardiology and two within the domain of pulmonology. Their task was to study, diagnose, recall, and explain the signs and symptoms of all the clinical case descriptions. The results indicate that cardiologists and pulmonologists alike, diagnosed cases within their specialty faster and more accurately. However, recall and pathophysiological explanations did not reveal significant differences. These results suggest that medical specialists do not process cases outside their specialty in a qualitatively different mode as compared with cases within their specialty.

¹ Parts of this chapter appeared as Rikers, R.M.J.P., Boshuizen, H.P.A., & Schmidt, H.G. (1997, March). *Nonroutine problem solving by medical experts*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, ILL.

Introduction

Theories on the development of expertise suggest that experts, while solving problems within their domain of expertise, apply knowledge that is qualitatively different from that applied by intermediates or novices. Koedinger and Anderson (1990), for instance, demonstrated, that in the domain of geometry problem-solving, experts do not rely on theorem proving rules and algebra, which is typical for participants of an intermediate level of expertise, but have developed prototypical schemas. These so-called *diagram configuration schemas* consist of clusters of geometric facts associated with a single prototypical geometric image, and are activated by relevant cues in the problem. The use of these schemas may explain why experts, in contrast to participants of lower levels of expertise, appear to skip steps while proving theorems (cf. Blessing & Anderson, 1996). This shift in organization of knowledge does not merely reflect a form of *chunking* or the formation of *macro-operators* (Kintsch & Van Dijk, 1978), but rather the emergence of qualitatively different structures more adapted to the task at hand (Koedinger & Anderson, 1990). A similar shift has been reported by Elio and Scharf (1990) who modeled the solution of physics word problems. In their view expert problem solving essentially consists of searching for relevant problem schemas and matching these schemas to features of the problem. By contrast, novices and advanced students appear to use a more bottom-up procedure, because they do not have appropriate schemas available for comprehending and solving problems of the domain.

Within the medical domain similar findings were reported by Feltovich and Barrows (1984) who suggested that in the course of practice, physicians develop cognitive structures of various diseases which they called *illness scripts*. These illness scripts guide a physician through a case, and comprise the idiosyncratic knowledge of disease and its consequences, in addition to knowledge of the situation in which a disease will likely occur. In essence, an illness script is a stereotyped sequential schema with three components: Enabling conditions, Fault, and Consequences. The first component, the Enabling Conditions, refers to the conditions under which a disease will likely occur. These are personal, social, medical, hereditary and environmental factors that affect the course of a specific disease. The second component, the Fault, refers to the (patho) physiological processes taking place in a specific disease. Finally, the third component consists of the Consequences of the Fault: The signs and symptoms related to a specific disease. While solving a problem, a physician activates one or a few illness scripts and they are matched with the information provided by the patient. In the course of this verification process the script becomes instantiated, i.e. expected values are substituted by actual findings. For an experienced physician this is an automatic and unconscious process as long as the information matches the activated illness script, and therefore no active reasoning is required (Custers, Boshuizen, & Schmidt, 1996; Hobus, Schmidt, Boshuizen, & Patel, 1987).

A medical student, on the other hand, due to a lack of experience, does not have (adequate) scripts available to construct a clinical case representation. According to Patel, Evans and Groen (1989), medical students generally construct a representation of a clinical case using primarily biomedical knowledge. That is,

knowledge concerning the mechanisms or processes underlying the manifestations of disease. Biomedical knowledge, which is taught in the first years in medical school, is helpful when it comes to comprehending general pathophysiological processes underlying signs and symptoms, but is not in itself a sufficient condition for constructing a complete and integrated clinical template belonging to a certain disease or category of diseases (Feltovich, Spiro, & Coulson, 1989). However, as competence is attained, elements of biomedical knowledge become increasingly interconnected, so that they can be accessed as coherent chunks of information rather than fragments. At this stage clinical reasoning processes hardly involve any biomedical concepts. Therefore, experienced physicians predominantly use, as a result of confrontations with real patients (instead of paper cases), clinical knowledge to construct a clinical case representation. Clinical knowledge is defined here as knowledge of attributes of sick people. It concerns the ways in which a disease can manifest itself in patients; the kind of complaints one would expect etc.

Similar results were found by Schmidt and Boshuizen (1992, 1993a) in a series of studies on clinical case representations by novices, advanced students and expert physicians. They found that *intermediates* (i.e., advanced medical students), while diagnosing clinical cases, applied large numbers of biomedical propositions in order to account for the clinical symptoms displayed in the cases. Experienced physicians, however, only applied biomedical knowledge to a limited extent. Schmidt and Boshuizen explained this difference between intermediates and expert physicians by assuming that a radical change has taken place in the expert's knowledge structure. Novices work on the basis of the surface features of a problem, make inferences and identify principles that subsume its surface structure. Under the influence of extensive practice and the confrontation with many real patients, the expert's biomedical knowledge structure has become packaged or *encapsulated* under a limited number of clinically relevant concepts that have the same explanatory power as the original biomedical structure. This restructuring, which has taken place in the course of the development from novice to expert, leads eventually to abbreviations in the lines of reasoning. For instance, a cardiologist and an advanced medical student are required to study a clinical case description that contains, among other things, the following information: *A 45-year old woman who has always felt a fast heartbeat. Her weight was 50 kg., and her height 1.70 m. Her thyroid gland was enlarged and elastic on palpation* (See Appendix 2.1 for a complete description of this case). The cardiologist would tend to process this description as "a condition of hyperthyroidism." This is a high-level inference (i.e., based on more than one proposition in the text) from the text that summarizes or encapsulates the information that is given. The cardiologist processes this information in an encapsulated mode, suggesting that this set of individual items in the text is processed as an integrated whole. In other words, an expert physician makes inferences and identifies principles that subsume the surface structure of a problem. An advanced student, on the other hand, will likely process the information as it is given, without the high-level inferences as made by the cardiologist. Obviously, this does not necessarily mean that the student eventually will not solve the patient's problem. Rather, it suggests that there is an important distinction between the representation created by the student and the one created by the cardiologist based on the same clinical case description (for a more

detailed discussion, see Schmidt & Boshuizen, 1993a). This dissociation of expertise level and memory performance was not only demonstrated within the medical domain. The most striking example is provided by a study of computer programmers by Adelson (1984). She found that novices actually had better memory for details of code than did experts. The reason appeared to be that experts attended more to the overall goal structure of the programming task, rather than the actual code. The experts found it easier to solve a programming task again rather than to memorize a detailed solution, whereas the reverse was the case for novices.

The difference between medical students and expert physicians in the way a clinical case is processed does account for the so-called *intermediate effect* in clinical case representation studies (e.g., Boshuizen, 1989; Claessen & Boshuizen, 1985; Schmidt & Boshuizen, 1993a; Schmidt, Boshuizen & Hobus, 1988). In these studies participants of different levels of expertise were requested to recall everything they could remember from a clinical case description. The outcome of these studies typically was that participants of an intermediate level of expertise produced a more elaborate recall than either experts or novices. Furthermore, the intermediate effect related to the level of expertise was not only found in recall studies, but this effect was also found in think-aloud and pathophysiology data as well. Experts who were asked to think aloud while diagnosing a clinical case, or explain the described signs and symptoms, applied more encapsulating concepts than either intermediates or novices (Boshuizen, Schmidt, & Coughlin, 1987; Schmidt & Boshuizen, 1993a).

Schmidt and Boshuizen (1993a) have demonstrated that the intermediate effect appears to be dependent on the amount of time available for processing a clinical case description. In their study the participants were required to study a case for either 3.30 min., 1.15 min., or 30 sec. The text was removed, and the participants had to reproduce everything they could remember from the text. What Schmidt and Boshuizen found, were only minor, nonsignificant differences in the amount of recall of the expert group (internists) between the different study times. The performance of the students, on the other hand, was severely and significantly impaired if study times were reduced. As matter of fact, the intermediate effect completely disappeared if processing time was restricted to 30 seconds, and the well-known increase in recall performance as a function of expertise emerged. These findings confirm the hypothesis that by reducing processing time, students do not have sufficient time to engage in deeper pathophysiological reasoning and so lost their advantage in terms of text recall. Expert physicians, however, are still able to construct a representation, because their representation is built upon a small number of highly inclusive encapsulating concepts and, therefore, they do not have to engage in deep case processing (Schmidt & Boshuizen, 1993a; Gilhooly, McGeorge, Hunter, Rawles, Kirby, Green, & Wynn, 1997).

In a series of studies by Patel, Arocha, and Groen, expert physicians from distinct medical specializations were compared on tasks within and outside their domains of expertise (Patel, Arocha, & Groen, 1986; Patel, Groen, & Arocha, 1990). They were instructed to read a text describing a patient's history, complaints and some additional findings such as physical examination and laboratory data. Regardless of their specialty, all physicians, that is, subexperts (i.e., specialists confronted with clinical cases outside their specialty) and experts alike, were able to recall the

information presented in the clinical cases. Furthermore, Patel and her colleagues have demonstrated that practitioners diagnosing a clinical case in their own domain of expertise focused on the clinical aspects of the problem. However, when they had to diagnose a clinical case outside their domain of expertise, their focus shifted towards the biomedical aspects of the problem (Patel & Groen, 1986, 1991).

Schmidt and Boshuizen assume that biomedical knowledge, although not used by expert physicians solving problems within their specialty, remains easily accessible when the need arises. For instance, when dealing with very complicated or novel problems. They have postulated that the different knowledge structures acquired by expert physicians in the course of the development towards expertise are still part of their knowledge base, and are accessed when more recently acquired structures fail to produce an adequate representation of a clinical problem (Schmidt & Boshuizen, 1993a; Schmidt, Boshuizen, & Norman, 1992). Furthermore, Schmidt and Boshuizen assume that the application of encapsulated knowledge is confined to the specialist's domain of expertise (Schmidt & Boshuizen, 1993b). Therefore, if expert physicians have to deal with clinical cases outside their specialty, they will not be able to transfer their encapsulated knowledge and, therefore, relapse into the biomedical processing approach. This implies, as most current theories of expertise, a rather extreme domain specificity (for a review, see Gick and Holyoak, 1987). On the other hand, expert physicians confronted with a clinical case outside their specialty could perform as well as within their specialty, because the case reminds them, in an abstract way, of similar problems previously solved (Schraagen, 1993). Practically, finding evidence for positive transfer of expert knowledge to novel situations could have educational implications: The representations used by experts could be made explicit and perhaps successfully taught to novices (Bereiter & Scardamalia, 1986; Palincsar & Brown, 1984; Schoenfeld, 1985).

The present study investigated the role of encapsulated knowledge as a general characteristic of medical expertise. Two groups of medical specialists (cardiologists and pulmonologists) diagnosed four clinical cases of which two were within, and two were outside their domain of expertise. Thus, both the cardiologists and pulmonologists fulfilled the role of expert and subexpert in this study. Based on the assumptions that (a) expertise development in a domain is characterized by the development of a rich and strongly linked knowledge base. That is, biomedical knowledge has become encapsulated under more clinically relevant concepts and (b) that this process of encapsulation accounts for the increase in processing speed and diagnostic accuracy as observed in the performance of medical specialists. It is predicted that medical specialists will relapse into elaborate biomedical case processing and are not able to utilize their encapsulated knowledge, if they are confronted with cases outside their specialty. Therefore, it can be predicted that, if medical specialists have to solve a clinical problem within their domain of expertise, they will mainly tend to rely on encapsulated knowledge. If, on the other hand, the problem is relatively unfamiliar, relevant encapsulating concepts will not have been developed and they have to resort to biomedical knowledge, which they have acquired earlier, in order to construct, a suitable problem representation (much like intermediates do).

Method

Participants

Participants were 16 experienced medical specialists (8 cardiologists and 8 pulmonologists) employed at Maastricht University Hospital, and from three other hospitals in the southeastern part of the Netherlands. The cardiologists and pulmonologists were practitioners with at least a MD degree and Board Certification in their specialty. All participants were practicing physicians and had on average more than 8 years of experience in their fields. The participants received a financial compensation for their participation.

Material

The materials consisted of four clinical case descriptions presented on a computer screen. The clinical cases were constructed by two domain experts and consisted of two cardiological and two pulmonary cases. The cases were based on actual patients treated at the Maastricht University Hospital. Each clinical case description reported some contextual information, the complaints, findings from history taking and physical examination, relevant laboratory data and some additional findings. Two cases were within the domain of cardiology (a case of Aortic valve insufficiency and Congestive heart failure) and two were pulmonary cases (a case of Small cell lung cancer and Legionnaires' disease). The translated text of one of the cases (Congestive heart failure) is provided in Appendix 2.1. The domain problems for one group of specialists were used as the non-domain cases for the other group and vice versa. The case descriptions were about one page in length and consisted of 94, 80, 104, and 105 propositions respectively.

Procedure

The basic experimental procedure was essentially the same as the one used by Patel and Groen (1986), Patel, Groen, and Arocha (1990), and Schmidt and Boshuizen (1993a), with the distinction that there were no restrictions on reading time. The participants were instructed to read each case carefully. After they had completed this task, they were asked to provide a diagnosis for the patient in question. Then they had to respond to a short question on a 5-point scale that enabled them to indicate how confident they were about the correctness of their diagnosis (1 = very insecure, 5 = very secure). After they had completed these tasks, a free recall task followed. The participants had to write down everything they could remember about the presented case. The participants knew beforehand that each case would be followed by a recall task. Finally, they had to explain, again in writing, the pathophysiology underlying the signs and symptoms of the case.

The complete texts of the clinical cases were divided into segments, and these were presented sequentially on a computer screen (cf. Joseph & Patel, 1990; Bordage & Lemieux, 1991; Lemieux & Bordage, 1992). New segments appeared on the screen by pressing the mouse button. Reading times were registered automatically after the

button was pressed. It was impossible to read a presented segment again after the mouse button had been pressed. Each participant had to evaluate all the cases. Carry-over effects were controlled by randomizing the presentation order of the cases. The participants were free to spend as much time as they needed to complete the different assignments. To enable the participants to get used to their task, a sample case (within the domain of gynaecology) preceded the four experimental cases.

Analysis

For each case the accuracy and completeness of the diagnosis were independently rated by two domain experts on a 7-point scale (from 0 = completely incorrect to 6 = completely correct diagnosis). If the accurate diagnosis was given as a part of a differential diagnosis it was considered accurate.

The recall protocols and the pathophysiological explanations were segmented into propositions using a technique from Kintsch (1974). Propositional analysis consists of segmenting a clinical case into individual propositions corresponding to discrete idea units in the text. The propositions consisted of two concepts connected by a qualifier, such as causation, negation, location and specification. For each proposition it was decided whether or not it matched any proposition in the text and repetitions were removed. The recall score was based on the total number of matching propositions. Evidence for encapsulation of signs and symptoms was explored by counting the number of high-level inferences in the recall protocols (Coughlin & Patel, 1986). High-level inferences were included in the analysis to the extent that they could be matched to a combination of propositions in the original text. Inferences referring to only one proposition in the text were excluded, because these could not be considered encapsulations. To distinguish high-level inferences from inferences based on only one proposition, they will be further referred to as *summaries*.

Each pathophysiological explanation was translated into a conceptual network that formalized the semantic relations underlying the disease (See for a more detailed discussion, Patel & Groen, 1986). Four measures were used to score the pathophysiology protocols. First, the elaborateness of their explanations was looked at by counting the number of different concepts they used. Second, the total number of links between these concepts was counted. The last two measures were derived from the comparison of the participants' explanation with a so-called *canonical explanation* of the signs and symptoms (Patel & Groen, 1986). A canonical explanation is constructed with the help of domain experts and provides a standard explanation of the signs and symptoms described in a case. The number of concepts in the participants' explanation identical with, or equivalent to the canonical model was counted. These concepts were termed *model concepts*. Finally, the number of links between concepts in the participants' explanations, that were identical to the links in the model, was also counted (the *model links*). Inter-rater agreement for each of these procedures was higher than 95 %.

The data were analyzed with a 2 x 2 analysis of variance, with Specialization as the between-groups factor and Case Type as the within-groups factor. The factor case

type refers to the mean score on the two cardiological cases and the mean score on the two pulmonary cases.

Results

Case reading times

Analysis of variance of the reading times, with Specialization as between-groups factor and Case Type as within-groups factor, showed a nonsignificant main effect of Specialization ($F < 1$), and Case Type, $F(1,14) = 2.58, p > .1$, but there was a significant interaction between Specialization and Case Type, $F(1,14) = 26.55, p < .0001$.

As is shown in Table 2.1, the cardiologists were on average faster on their own cases than the pulmonologists who were faster on the pulmonary cases, which were considered domain problems for them.

Table 2.1
Mean Reading Times in seconds as a Function of Case Type and Specialization

Case Type	Cardiologists		Pulmonologists	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Cardiology	135.65	60.31	166.48	56.19
Pulmonology	166.11	61.00	150.50	59.31

Diagnostic accuracy

Table 2.2 shows the main accuracy of the diagnoses provided by the physicians as a function of Specialization and Case Type. No significant differences between cardiologists and pulmonologists were found ($F < 1$), but there was a significant main effect of Case Type, $F(1,14) = 33.91, p < .0001$, and a significant interaction, $F(1,14) = 16.97, p < .001$. Consequently, both cardiologists and pulmonologists surpassed each other within their domains of expertise on the measure of diagnostic accuracy. Furthermore, both groups diagnosed the pulmonary cases better than the cardiological cases. This finding indicates that the pulmonary cases were less complicated to diagnose than the cardiological cases.

Table 2.2
Mean Accuracy of Diagnoses as a Function of Case Type and Specialization

Case Type	Cardiologists		Pulmonologists	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Cardiology	3.69	0.53	2.75	0.96
Pulmonology	4.06	0.86	4.94	0.68

The diagnoses were accompanied by a confidence rating. The participants had to indicate how confident they were that the diagnosis provided was completely

correct on a 5-point scale. The analysis of the confidence rating showed no significant main effect of Specialization ($F < 1$), but there was a significant main effect of Case Type, $F(1,14) = 5.50, p < .05$, and a significant interaction, $F(1,14) = 44.57, p < .0001$. Expert physicians were more confident about their diagnosis than the subexperts (Table 2.3). The measures of diagnostic accuracy and confidence should be considered as treatment checks, as to what extent the cases presented could be considered familiar or unfamiliar problems for the respective participants.

Table 2.3
Mean Confidence Ratings as a Function of Case Type and Specialization

Case Type	Cardiologists		Pulmonologists	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Cardiology	4.63	0.44	3.56	0.50
Pulmonology	3.06	0.50	4.31	0.37

Free Recall

Table 2.4 shows the mean number of propositions recalled. There was no significant main effect of Specialization ($F < 1$). The main effect of Case Type was significant, $F(1,14) = 13.73, p < .01$. This finding can largely be explained if one would take the differences in length (i.e., number of propositions) between the cases into consideration: The cardiological cases were both shorter than the pulmonary cases. The interaction between Specialization and Case Type was not significant, $F(1,14) = 2.77, p > .1$.

Table 2.4
Mean Numbers of Propositions Recalled as a Function of Case Type and Specialization

Case Type	Cardiologists		Pulmonologists	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Cardiology	24.06	10.31	22.50	9.31
Pulmonology	26.25	10.10	28.25	10.73

Further, the presence of high-level inferences or summaries in the recall protocols was looked at, because they are considered as important indicators whether or not encapsulating concepts had been activated (Table 2.5). The analysis showed no significant main effects of Specialization, $F(1,14) = 1.33, p > .25$, and Case Type ($F < 1$) and there was no interaction ($F < 1$). Therefore, medical experts did not introduce more encapsulating concepts as compared with subexperts.

Table 2.5
Mean Numbers of Summaries as a Function of Case Type and Specialization

Case Type	Cardiologists		Pulmonologists	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Cardiology	3.13	1.38	2.62	1.09
Pulmonology	3.31	2.12	2.56	1.18

Pathophysiological explanations

It was assumed that experts, while explaining clinical problems within their specialty, apply qualitatively different knowledge compared with problems outside their specialty. Four measures were used in order to analyze the structure of the participants' pathophysiological protocols.

Table 2.6
Mean Numbers of Concepts as a Function of Case Type and Specialization

Case Type	Cardiologists		Pulmonologists	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Cardiology	6.69	1.78	7.19	2.52
Pulmonology	6.69	3.27	7.56	2.23

Table 2.6 shows the total number of different concepts the experts used to explain the signs and symptoms described in the cases. The analysis showed no significant main effects of Specialization and Case Type, and there was no interaction ($F_s < 1$). Table 2.7 shows the total number of links between concepts in their explanation. Again, there were no significant main effects of Specialization and Case Type, and there was no interaction ($F_s < 1$).

Table 2.7
Mean Numbers of Links as a Function of Case Type and Specialization

Case Type	Cardiologists		Pulmonologists	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Cardiology	6.00	2.05	5.50	2.95
Pulmonology	5.31	2.89	6.00	2.19

The last two measures to be reported here are based on the overlap of the physicians' explanations with the canonical models (Table 2.8). First, the number of concepts in the participants' explanation which were identical or equivalent to concepts in the canonical explanation were computed (the model concepts). The analysis did not show significant main effects of Specialization and Case Type, and there was no interaction ($F_s < 1$).

Table 2.8
Mean Numbers of Model Concepts as a Function of Case Type and Specialization

Case Type	Cardiologists		Pulmonologists	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Cardiology	4.25	2.02	4.50	2.02
Pulmonology	4.56	1.71	5.31	1.71

Finally, the number of links between concepts in the participants' explanations that were identical to the links in the model was analyzed, the so-called model links (Table

2.9). There were no significant main effects of Specialization and Case Type and there was no interaction ($F_s < 1$).

Table 2.9

Mean Numbers of Model Links as a Function of Case Type and Specialization

Case Type	Cardiologists		Pulmonologists	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Cardiology	1.44	0.73	0.38	0.44
Pulmonology	2.13	1.79	1.00	0.35

Discussion

Clinical case studies have repeatedly demonstrated the pivotal role of encapsulating knowledge within the physician's domain of expertise (e.g., Boshuizen, 1989; Schmidt & Boshuizen, 1993a; Schmidt, Boshuizen, & Norman, 1992). Furthermore, the intermediate effect, as it was found in clinical case studies comparing the performance of expert physicians with that of students, is probably the most striking exponent of the role of encapsulated knowledge. In order to disentangle the issue of the role of encapsulated knowledge outside the physicians' domain of expertise, the present study compared the performance of medical experts with that of subexperts.

It was predicted that medical experts, as a direct result of their encapsulated knowledge, will solve clinical cases within their domain of expertise, as compared with subexperts, faster and more accurately, and that their recall and pathophysiological explanations would be more compendious, because they contain many encapsulating concepts. Medical specialists were expected to fail in activating relevant encapsulating concepts outside their specialty, because encapsulated knowledge was expected to be confined to the specialist's domain of expertise.

The present study partially corroborated these predictions. Reading times, diagnostic accuracy and confidence were in accordance with the expectations: medical experts solved the clinical cases faster, more accurately, and they are more confident about their diagnoses than subexperts. Cardiologists were approximately 23% faster at diagnosing cases within their specialty than outside and pulmonologists were approximately 11% faster. Although cardiologists did not diagnose their own cases more accurately than pulmonary cases, the interaction effect for diagnostic accuracy, shows that indeed diagnostic accuracy is a function of specialization. These findings do indicate that medical experts construct a representation of a clinical case much more rapidly and more accurately than medical subexperts do. Therefore, the question that has to be answered is, why are experts faster and more accurate than subexperts? Our preliminary answer to this question was that experts process cases within their domain of expertise using primarily encapsulating knowledge, enabling them to process information in bigger, and qualitatively different, chunks, as compared with subexperts. This assumption was based on findings in previous clinical case studies comparing experts' performance with that of advanced students (e.g., Boshuizen, 1989; Claessen & Boshuizen, 1985; Schmidt & Boshuizen, 1993a; Schmidt, Boshuizen & Hobus, 1988; Simpson & Gilhooly, 1997). These stud-

ies have demonstrated that experts produce more condensed recall as compared with intermediates, and recall protocols contain more summaries. In addition, experts' pathophysiological explanations have been shown to contain more encapsulating concepts, both in on-line and post-hoc explanation tasks. These findings were taken as evidence for encapsulated processing of clinical case information by experts.

The present study demonstrated no differences in recall and pathophysiological measures between medical experts and subexperts. The total amount of recall produced did not show an interaction effect favoring the subexperts. In addition, number of summaries produced in the recall did not significantly differ between experts and subexperts. The pattern unfolded in the analysis of the total amount of recall and summaries, concerning the important interaction effect, seems to continue in the pathophysiological explanations. No significant interaction between Case Type and Specialization was found for number of concepts used, number of links between concepts used, and the measures of overlap with the canonical explanations. The assumption that cases outside their specialty would evoke more elaborate pathophysiological explanations, because of the absence of adequate encapsulated concepts was therefore not corroborated.

The question, then, is how these findings must be interpreted. A first, and fairly straightforward explanation would be that medical specialists process cases within and outside their domain in essentially the same way. Qualitative differences in clinical case processing as found between experts and intermediates do not apply when comparing experts with subexperts. Experts and subexperts essentially process clinical cases in similar fashion; experts only do it faster and more accurately, presumably because they have seen more cases similar to those presented. This is, however, not an entirely satisfying hypothesis, because it leaves unexplained how frequent exposure to similar problems results in an increase of accuracy and processing speed.

A second possible explanation is based on the fact that both groups included in the experiment did not extremely well on the measure of diagnostic accuracy. The maximum score for a completely correct diagnosis was six points, and there were four opportunities for each participant to attain this score: two within and two outside their specialty. Just five pulmonologists and only two cardiologists attained the maximum score (all on the case of legionnaires' disease). So, of the thirty-two diagnoses provided by the cardiologists together only two (6.25%) were completely correct and only five (15.63%) of all diagnoses provided by the pulmonologists. This might indicate that the clinical cases used within and outside the specialists' domain of expertise were complex problems, and they were not able to process these cases routinely. As a result, the medical specialists might have relapsed into an elaborate processing approach, which is typical for students of an intermediate level of expertise.

A third reason why the present study did not unearth differences with regard to indicators of encapsulation, may be that the two domains studied are more or less adjacent. In the Netherlands, cardiologists and pulmonologists share the first two years of their training during their residency. This is a period where they have to evaluate cardiological and pulmonary problems on a regular basis. Possibly, encapsulation of biomedical knowledge takes place quite early in residency training,

resulting in similar levels of encapsulated knowledge for both specializations. It is also to be noted that the present study replicates findings of Patel et al. (1986, 1990) who found no differences in recall performance either, although they did use experts of nonadjacent domains. Patel and her co-workers assumed that medical experts and subexperts have developed some kind of *generic expertise* for the relevant information in clinical cases. Generic expertise implies the existence of 'some schema or macrostructure representation (Fletcher, & Chrysler, 1990; Kintsch, 1988; Kintsch & Van Dijk, 1978) that is based on their experience in practical problems, and guarantees the retention of crucial facts in a clinical case within or outside their domain of expertise. Therefore, subexperts are able to recall and explain the essential information of a clinical case. Domain experts, on the other hand, also possess the specific expertise, which is crucial in order to achieve a high diagnostic accuracy. They possess the necessary domain knowledge that enables them to link the non-salient cues or loose ends to the main diagnosis (Patel & Groen, 1991; Patel, Groen, & Arocha, 1990). This explanation seems plausible, especially in this case where the two domains of expertise were related. Moreover, it is important to note that it is almost impossible to compare physicians on the pathophysiological measures having specializations that are very remote, because there is a great risk that these physicians are not able to evaluate each other's cases at all. After all, when one assumes that experts produce a more elaborate pathophysiological explanation outside their specialty, one also implicitly assumes that they at least possess the basic pathophysiological knowledge of the domain.

In sum, although expert physicians do not seem to apply different forms of knowledge to deal with their diagnostic task, as compared with subexperts, their performance on the diagnostic, speed and confidence measure differs significantly from that of subexperts. Furthermore, this study seems to suggest that the performance of expert physicians on the measures of recall and pathophysiology are relatively insensitive for the nature of the cases used (i.e., within or outside their specialty). It is, therefore, not unlikely that the intermediate effect as it was originally found in clinical case studies comparing expert physicians with medical students, will also emerge if the expert physicians are confronted with clinical cases outside their specialty. Further research is needed to investigate the role of encapsulated knowledge outside the specialists' domain of expertise, and its contribution to the intermediate effect.

References

- Adelson, B. (1984). When novices surpass experts: The difficulty of a task may increase with expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 483-495.
- Bereiter, C., & Scardamalia, M. (1986). Educational relevance of the study of expertise. *Interchange*, 17(2), 10-19.
- Blessing, S. B., & Anderson, J. R. (1996). How do people learn to skip steps? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22 (3), 576-598.
- Bordage, G. and Lemieux, M. (1991). Semantic structures and diagnostic thinking of experts and novices. *Academic Medicine*, 66, S70-S72.

- Boshuizen, H. P. A. (1989). *De ontwikkeling van medische expertise; een cognitief-psychologische benadering*. [On the development of medical expertise; a cognitive psychological approach]. Doctoral dissertation. Haarlem: Thesis.
- Boshuizen, H. P. A., Schmidt, H. G., & Coughlin, L. D. (1987). *On-line representation of a clinical case and the development of expertise*. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, DC.
- Claessen, H. F. A., & Boshuizen, H. P. A. (1985). Recall of medical information by students and doctors. *Medical Education*, 19, 61-67.
- Coughlin, L. D., & Patel, V. L. (1986, April). *Text comprehension and expertise in the domain of medicine*. Paper presented at the Meeting of the American Educational Research Association, San Francisco, CA.
- Custers, E. J. F. M., Boshuizen, H. P. A., & Schmidt, H. G. (1996). The influence of medical expertise, case typicality, and illness script component on case processing and disease probability estimates. *Memory and Cognition*, 24(3), 384-399.
- Elio, R., & Scharf, P. B. (1990). Modeling novice-to-expert shifts in problems-solving strategy and knowledge organization. *Cognitive Science*, 14, 579-639.
- Feltovich, P.J. & Barrows, H.S. (1984). Issues of generality in medical problem solving. In: Schmidt, H.G. & De Volder, M.L. (Eds.), *Tutorials in problem-based learning* (pp. 128-142). Assen: Van Gorcum.
- Feltovich, P. J., Spiro, R. J., & Coulson, R. L. (1989). The nature of conceptual understanding in biomedicine: The deep structure of complex ideas and the development of misconceptions. In D. A. Evans & V. L. Patel (Eds.), *Cognitive science in medicine: Biomedical modeling*. Cambridge, MA: MIT Press.
- Fletcher, C. R., & Chrysler, S. T. (1990). Surface forms, text bases, and situation models: Recognition memory for three types of textual information. *Discourse Processes*, 13(2), 175-190.
- Gick, M.L., & Holyoak, K.J. (1987). The cognitive basis of knowledge transfer. In S.M. Cormier & J.D. Hagman (Eds.), *Transfer of training: Contemporary research and applications* (pp. 9-46). New York: Academic Press.
- Gilhooly, K. J., McGeorge, P., Rawles, J. M., Kirby, I. K., & et al. (1997). Biomedical knowledge in diagnostic thinking: The case of electrocardiogram (ECG) interpretation. *European Journal of Cognitive Psychology*, 9(2), 199-223.
- Hobus, P. P., Schmidt, H. G., Boshuizen, H. P., & Patel, V. L. (1987). Contextual factors in the activation of first diagnostic hypotheses: Expert-novice differences. *Medical Education*, 21(6), 471-476.
- Joseph, G. M., & Patel, V. L. (1990). Domain knowledge and hypothesis generation in diagnostic reasoning. *Journal of Medical Decision Making*, 10, 31-46.
- Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, NJ: Erlbaum.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95(2), 163-182.
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363-394.
- Koedinger, K. R., & Anderson, J. R. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science*, 14, 511-550.
- Lemieux, M., & Bordage, G. (1992). Propositional versus structural semantic analyses of medical diagnostic thinking. *Cognitive Science*, 16, 185-204.
- Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1(2), 117-175.

- Patel, V. L., Arocha, J. F., & Groen, G. J. (1986). Strategy selection and degree of expertise in medical reasoning. In C. Clifton (Ed.), *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*, (pp. 780-791). Hillsdale, NJ: Erlbaum.
- Patel, V. L., Evans, D. A., & Groen, G. J. (1989). Biomedical knowledge and clinical reasoning. In D. A. Evans & V. L. Patel (Eds.), *Cognitive science in medicine: Biomedical modeling*, (pp. 53-112). Cambridge, MA: MIT Press.
- Patel, V. L., & Groen, G. J. (1986). Knowledge based solution strategies in medical reasoning. *Cognitive Science*, 10, 91-116.
- Patel, V. L., & Groen, G. J. (1991). The general and specific nature of medical expertise: A critical look. In A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits*, (pp. 93-125). New York: Cambridge University Press.
- Patel, V. L., Groen, G. J., & Arocha, J. F. (1990). Medical expertise as a function of task difficulty. *Memory and Cognition*, 18(4), 394-406.
- Schmidt, H. G., & Boshuizen, H. P. A. (1992). Encapsulation of biomedical knowledge. In D. A. Evans & V. L. Patel (Eds.), *Advanced models of cognition for medical training and practice*. New York, NY: Springer Verlag.
- Schmidt, H. G., & Boshuizen, H. P. A. (1993a). On the origin of intermediate effects in clinical case recall. *Memory and Cognition*, 21(3), 338-351.
- Schmidt, H. G., & Boshuizen, H. P. A. (1993b). On acquiring expertise in medicine. *Educational Psychology Review*, 5(3), 205-221.
- Schmidt, H. G., Boshuizen, H. P. A., & Hobus, P. P. M. (1988). Transitory stages in the development of medical expertise: The "intermediate effect" in clinical case representation studies, *Proceedings of the Cognitive Science Society Meeting*, (pp. 139-145). Hillsdale, NJ: Lawrence Erlbaum.
- Schmidt, H. G., Boshuizen, H. P. A., & Norman, G. R. (1992). Reflections on the nature of expertise in medicine. In E. Keravnou (Ed.), *Deep models for medical knowledge engineering*. Amsterdam, The Netherlands: Elsevier.
- Schoenfeld, A. H. (1985). Making sense of "out loud" problem-solving protocols. *Journal of Mathematical Behavior*, 4(2), 171-191.
- Schraagen, J. M. (1993). How experts solve a novel problem in experimental design. *Cognitive Science*, 17, 285-309.
- Simpson, S. A., & Gilhooly, K. J. (1997). Diagnostic thinking processes: Evidence from a constructive interaction study of electrocardiogram (ECG) interpretation. *Applied Cognitive Psychology*, 11(6), 543-554.

Appendix 2.1

Case of heart failure

A 45-year old woman has complaints of increasing dyspnea and ankle edema since six weeks. The patient has always felt a fast heartbeat.

History: Epileptic from the age of fifteen. Family history: Negative. Social history: She owns a homeopathic shop. Intoxications: She smokes 20 cigarettes a day. Medication: Depakine 3x300 mg.

Physical examination:

Physical examination revealed a thin, dyspnoeic woman. Her weight was 50 kg., and her height 1.70 m. Pulse rate was 100/min and regular. Blood pressure was 140/90 mmHg. Carotid artery pulses normal. Distended jugular veins. Thyroid gland was enlarged and elastic on palpation. Heart: Apical impulse was just lateral to the midclavicular line. First and second heart sounds are normal. Left and right ventricular pulsations. There was a 3/6 holosystolic murmur at the apex. Lungs: Normal breath sounds, basal rales and dullness on percussion. Abdomen: The liver was enlarged (four fingers). Extremities: Pronounced bilateral ankle edema.

Laboratory findings:

ESR 6 mm/h (normal: <12 mm/h), Hb 9.1 mmol/l (normal: 7.4-9.6 mmol/l), Ht 0.42 l/l (normal: 0.36-0.46 l/l), WBC $6.8 \cdot 10^9/l$ (normal: $3.4-8.9 \cdot 10^9/l$), Platelet count $153 \cdot 10^9/l$ (normal: $150-350 \cdot 10^9/l$), PTT 31 sec (normal: 24-35 sec), Sodium 140 mmol/l (normal: 132-142 mmol/l), Potassium 5.25 mmol/l (normal: 3.6-5.0 mmol/l), Calcium 2.27 mmol/l (normal: 2.1-2.6 mmol/l), Urea 4.5 mmol/l (normal: 3.0-7.0 mmol/l), Creatinine $67 \mu\text{mol/l}$ (normal: $53-97 \mu\text{mol/l}$), Alkaline phosphatase 129 U/l (normal: 30-125 U/l), GGT 42 U/l (normal: < 50 U/l), ASAT 27 U/l (normal: < 35 U/l), ALAT 19 U/l (normal: < 35 U/l), CPK 86 U/l (normal: 40-200 U/l), LDH 381 U/l (normal: 200-450 U/l), Total protein 68 g/l (normal: 65-79 g/l), Albumin 36 g/l (normal: 40-60 g/l), Arterial blood gases: pH 7.41 (normal: 7.35-7.45), pCO_2 5.1 kPa (normal: 4.5-5.9 kPa), pO_2 10.5 kPa (normal: 8.7-13.1 kPa), O_2 saturation 94% (normal: 93-98 %).

Electrocardiogram:

Sinus rhythm, rate 90/min. Left atrial enlargement. Intermediate position of the electrical axis. Poor R wave progression in the left precordial leads. Abnormal ST-T segments.

Chest X-ray:

Cardiothoracic ratio > 0.5. Bilateral pleural effusion (right > left).

Echocardiography/Doppler:

Global hypokinesia of the dilated left ventricle and severe mitral insufficiency. Mitral valvularannulus enlarged 4 cm. The mitral valve leaflets are slightly thickened.

Chapter 3

On the Constraints of Encapsulated Knowledge: Clinical Case Representations by Medical Experts and Subexperts (Part 2)¹

This article is concerned with the role of so-called "encapsulated knowledge" (Schmidt & Boshuizen, 1992, 1993) in the process of diagnosing clinical cases within and outside the medical specialist's domain of expertise. Based on the theory of knowledge encapsulation it was predicted that subexperts (i.e., medical specialists diagnosing a clinical case outside their specialty) have difficulty to diagnose a clinical case in an encapsulated mode, and, therefore, would relapse into an elaborate biomedical processing approach in order to understand the described signs and symptoms. Cardiologists (the experts) and neurologists (the subexperts) were instructed to study two clinical case descriptions for a period of three minutes per case. After each case they were asked to provide a diagnosis, write down everything they remembered of the case, and finally, to explain the signs and symptoms displayed in the case. The results showed that the cardiologists achieved a higher diagnostic accuracy than did the neurologists. Furthermore, the cardiologists' recall and pathophysiological explanations contained more high-level inferences (i.e., encapsulations) than those of neurologists did. These outcomes are in line with the predictions made on the basis of the theory of knowledge encapsulation.

¹ Parts of this chapter appeared as Rikers, R.M.J.P., Schmidt, H.G., Boshuizen, H.P.A. (1999, April). *Recall of clinical case descriptions by medical experts and subexperts*. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal.

Introduction

The measure of free recall has become popular as a measure of expertise ever since De Groot's (1946, 1965) studies in chess. In his studies, chess players of different levels of expertise had to recall a meaningful chess position after viewing the position for just a few seconds. De Groot demonstrated that the number of pieces recalled was a linear function of level of expertise. These findings have been replicated in many other domains such as computer programming (McKeithen, Reitman, Rueter, & Hirtle, 1981), bridge (Charness, 1979), Go (Reitman, 1976), mental calculation (Ericsson, 1985; Jensen, 1990), and baseball (Spilich, Vesonder, & Voss, 1979). A recent review of free recall studies by Vincente and Wang (1998) has shown that there are at least 51 studies in 19 different domains demonstrating the superior memory performance of experts on meaningful stimuli.

However, in the domain of medicine findings have been reported that seem to imply that the relationship between level of expertise and amount of recall is not as straightforward as in the aforementioned domains. For instance, studies on clinical case representations by participants of different levels of expertise have repeatedly demonstrated the so-called *Intermediate Effect* (e.g., Boshuizen, 1989; Boshuizen & Schmidt, 1992; Coughlin & Patel, 1986; Patel & Medley-Mark, 1986; Schmidt & Boshuizen, 1993; Schmidt, Boshuizen, & Hobus, 1988). In these studies medical students and expert physicians (i.e., medical specialists) were instructed to read a text describing a patient's history, complaint(s) and some additional findings such as physical examination and laboratory data. The outcome of these studies typically was that participants of an intermediate level of expertise (i.e., advanced medical students) recalled the clinical case information better than did either more or less experienced participants. So, rather than a monotonically increasing function of recall with increasing expertise, an inverted U-shaped relation was found.

An explanation for this paradoxical Intermediate Effect is provided by Schmidt and Boshuizen (1992, 1993). They suggest that novices, intermediates and expert physicians apply qualitatively different kinds of knowledge to deal with their task. The novices' medical knowledge, according to Schmidt and Boshuizen, can be best characterised by a limited understanding of biomedical knowledge and a lay type of clinical knowledge.² Intermediates have acquired, through individual study and lectures, extensive biomedical knowledge, but as the novices they only have little or no knowledge of the manifestations of a disease in real patients. Their clinical reasoning process consists of chains of small steps based on detailed, biomedical concepts. The expert physicians' knowledge base, however, is different from that of novices and intermediates in that biomedical knowledge only plays a minor and implicit role in their diagnostic reasoning (Boshuizen & Schmidt, 1992; Gilhooly, 1996; Gilhooly & Simpson, 1992; Simpson & Gilhooly, 1997). As a result of *deliberate* practice (Ericsson, Krampe, & Tesch-Römer, 1993) and the confrontation with

² Biomedical knowledge, pertains to the (pathological) processes underlying the manifestations of disease, and it incorporates knowledge about domains such as biochemistry, microbiology, and physiology. Clinical knowledge, on the other hand, concerns the ways in which a disease can manifest itself in patients (Patel, Evans, & Groen, 1989; Patel & Kaufman, 1995; Schmidt, Boshuizen, & Norman, 1992).

real patients, expert physicians' biomedical concepts have become clustered together, enabling expert physicians to make shortcuts in their lines of reasoning, skipping intermediary concepts (cf. Blessing & Anderson, 1996; Koedinger & Anderson, 1990). Schmidt and Boshuizen termed this process of knowledge restructuring, which takes place in the course of the development from novice to expert physician, "knowledge encapsulation." Many of these concept clusters or encapsulating concepts have clinical or even diagnostic labels, such as *aorta-insufficiency*, *backward failure*, *cyanosis*, *endocarditis*, *hepatolienomegaly*, or *hyperhidrosis*.

The activation of encapsulating concepts is a process that takes place while the clinical case information is studied. In other words, information about signs and symptoms in a case, point towards, and therefore activate, specific pre-existing encapsulating concepts in long-term memory (cf. Ericsson & Kintsch, 1995). Knowledge encapsulation explains, therefore, why experts' recall is much more compendious than that of intermediates: Experts process clinical cases almost entirely in an encapsulated mode. For instance, in the Schmidt and Boshuizen study (1992), one internist recalled the endocarditis case (see Appendix 3.1) in the following fashion: "The patient is a young man with a high fever who presents a *septic syndrome*. This suggests *drug use*. He shows signs of *thromboemboli*, due to an *affected heart valve*. The *tachycardia* fits with an *associated aorta vitium*" [italics added]. If we compare this internist's recall with the original text in the Appendix 3.1, it should be noted that it almost completely consists of inferences. The ability of expert physicians to grab the gist of the information presented, rather than its surface structure, is an important manifestation of the encapsulation of the knowledge base. Therefore, intermediates' free recall is not really better than that of expert physicians, unless one wants to equate more with better (Patel & Groen, 1991b). It is important to note that it takes thousands of hours of *deliberate practice* (Ericsson, Krampe, & Tesch-Römer, 1993) and the confrontation with large numbers of real patients before a physician's knowledge base can be characterised as encapsulated. As is in many other domains, the highest level of expert performance is displayed by physicians with more than 10 years of experience (Boshuizen, 1989; Ericsson & Lehmann, 1996; Lesgold, 1984; Patel & Groen, 1991a).

Evidence for knowledge encapsulation is not only found in recall data, but also in post-hoc pathophysiological explanations (Boshuizen & Schmidt, 1992; Schmidt & Boshuizen, 1992, 1993). Expert physicians, who were asked to explain the signs and symptoms of a clinical case, primarily referred to clinically relevant concepts. They made direct links between patient findings and clinical concepts that have the status of hypotheses or diagnoses in their reasoning process. Advanced students, on the other hand, primarily used biomedical concepts to account for the signs and symptoms. Furthermore, Schmidt and Boshuizen (1993) have demonstrated that the occurrence of an intermediate effect in recall and pathophysiology data is dependent on the amount of time available for processing a clinical case. In their study participants of different levels of expertise were subdivided into three groups who had to study a clinical case description for 3.30 minutes, 1.15 min., or 30 seconds. The results showed clear intermediate effects for recall and pathophysiology data under the 3.30 and 1.15 minute's conditions. However, if processing time

was restricted to 30 seconds, the intermediate effect disappeared completely: An indication that medical students did not have sufficient time to engage in elaborate biomedical reasoning in order to understand the clinical case description, and their performance in terms of recall and pathophysiology deteriorated dramatically. Expert physicians' performance, on the other hand, was almost insensitive for manipulations of time. This suggests that they did not have to engage in deep pathophysiological reasoning, and mainly constructed a clinical case representation based on a small number of encapsulating concepts.

However, if needed, expert physicians can unfold their encapsulated knowledge, for instance when dealing with a very complicated or unfamiliar problem. Schmidt and Boshuizen have postulated that the different knowledge structures acquired in the course of the development towards expertise are stored in long-term memory, and are still accessible when more recently acquired, encapsulated structures, fail to produce an adequate representation of a clinical problem (Boshuizen & Schmidt, 1992; Schmidt & Boshuizen, 1993; Schmidt, Boshuizen, & Norman, 1992). Evidence to support this relapse of expert physicians towards biomedical knowledge is provided by Patel, Arocha, and Groen (1986, 1990). They compared the performance of board certified medical specialists from two specialties (cardiology and endocrinology) in diagnosing clinical problems in the domains of cardiology and endocrinology. The specialists had to diagnose and explain clinical cases within and outside their domain of expertise. As might be expected, diagnostic accuracy decreased as a transition was made from the level of expert to that of subexpert. More interestingly, however, was that specialists (i.e., practitioners explaining a problem in their own domain of expertise) focused primarily on patient-oriented components (i.e., clinical aspects) of the problem. But when they had to explain a clinical case outside their domain of expertise, their focus shifted towards biomedical components of the problem. This seems to indicate that the application of biomedical knowledge of expert physicians is associated with complicated or non-routine problems. On the other hand, Schmidt and Boshuizen have also shown that expert physicians, as a result of task characteristics, are able to process a clinical case in either an elaborate biomedical mode or in a compendious encapsulated mode (cf. Bordage, 1994). In one of their studies, participants of different levels of expertise were instructed to state as much as possible about a disease called endocarditis. It was shown that expert physicians generated considerably more information than any other group (Schmidt & Boshuizen, 1993). This finding suggests that expert physicians generally process clinical cases in an encapsulated mode, but if the need arises, they easily can revert to their biomedical knowledge to explain the signs and symptoms.

The present study investigates the role of encapsulating knowledge in diagnosing clinical cases within and outside physician's domain of expertise. It is assumed that if expert physicians are confronted with problems outside their domain of expertise, they will shift back or relapse into the elaborate biomedical processing approach that is typical for students of an intermediate level of expertise. Cardiologists (the experts) and neurologists (the subexperts) were required to diagnose two cardiological clinical cases. Based on the assumptions that the development of expertise in a domain is characterized by knowledge encapsulation, and that this

knowledge is confined to the physician's domain of expertise, it can be predicted that subexperts will recall *more* information and provide *more* elaborate pathophysiological explanations of a clinical case description than do (domain) experts. However, the proportion of encapsulating concepts within the experts' recall protocols will be higher than that of subexperts. Furthermore, because domain experts have encountered more cases similar to those selected than did subexperts, their diagnostic accuracy will also be higher.

Method

Participants

Participants were 20 medical specialists (10 cardiologists and 10 neurologists) of the Maastricht University Hospital and from four other hospitals in the southeastern part of the Netherlands. The cardiologists and neurologists were practitioners with at least a MD degree and Board Certification in their specialty and had at least 5 years of experience. The participants received a financial compensation for their participation.

Material

The materials consisted of two booklets, each containing a clinical case description and blank response sheets. The clinical cases consisted of two cardiological cases (Heart failure and Acute bacterial endocarditis) and were constructed by two domain experts. The case of bacterial endocarditis was a Dutch translation of the case used by Patel and Groen (1986). The cases were based on actual patients and were presented following a standard patient-presentation format. Each clinical case description reported some contextual information, the complaint(s), findings from history taking and physical examination, relevant laboratory data and some additional findings. The case descriptions were about half a page in length and consisted of 43 and 49 propositions respectively. The complete text Bacterial endocarditis is provided in Appendix 3.1.

Procedure

The participants were shown two written clinical case descriptions for a period of 3 minutes per case. After they had studied each clinical case, participants were instructed to provide a diagnosis. Subsequently, they were instructed to write down everything they could remember about the case description. Finally, they had to explain in writing the signs and symptoms displayed the case: The so-called pathophysiological explanation. The participants were free to spend as much time as they needed to complete these different assignments. Each participant was tested individually and had to diagnose both cases. In order to avoid carry-over effects the presentation order of the cases was counterbalanced. To enable the participants to get used to their tasks, the experimental cases were preceded by one sample case.

Analysis

The accuracy of the provided diagnosis for each case was rated by two independent domain experts on a 7-point scale from 0 (completely incorrect diagnosis) to 6 (completely correct diagnosis). If an accurate diagnosis was given as part of a differential list of diagnoses, it was considered accurate. The analysis of recall and pathophysiological explanations consisted of segmenting the protocols into propositions using a technique from Kintsch (1974, 1988). The essential element in propositional analysis is the segmentation of a clinical case into individual propositions corresponding to discrete idea units in the text. Each proposition consists of two concepts connected by a qualifier, such as *causation* (cau.), *negation* (neg.), *location* (loc.), or *specification* (spec.). For instance, the protocol fragment “The heart is not enlarged and auscultation reveals a 3/6 holosystolic murmur at the apex” consists of 5 propositions: 1. heart-neg. (enlarged); 2. auscultation-spec (murmur); 3. murmur-spec (3/6); 4. murmur-spec (holosystolic); 5. murmur-loc (at the apex). For each proposition in the free recall, it was decided whether or not it matched any proposition in the text. The identified propositions are divided into four types: Literal propositions, paraphrased, low-level and high-level inferred propositions. Low-level inferences are based on only one proposition in the text, and should, therefore, not be considered as encapsulations. Evidence for encapsulation of signs and symptoms was examined by counting the number of high-level inferences in the recall protocols. High-level inferences were included in the analysis to the extent that they could be matched to a combination of propositions in the original text (Coughlin & Patel, 1986). For instance, if the text contains the following information: *An old woman has always felt a fast heartbeat. Physical examination reveals a thin woman. Her weight is 50 kg. and her height 1.70 m. Pulse rate was 100/ min. and regular. Thyroid gland is enlarged and elastic on palpation.* A possible high-level inference would be a woman with a condition called *hyperthyroidism*, because this inference is based on all the information that is provided. To distinguish high-level inferences from inferences based on only one proposition, they will be further referred to as *summaries*. The total recall score was based on the number of literally recalled propositions, paraphrased propositions, low-level inferences, and summaries.

Four measures were used in order to analyse the pathophysiological explanations. First, the number of different concepts used in the participants’ explanations was counted. Second, the number of links between the concepts was counted. The last two measures were based on the overlap with a so-called *canonical explanation* of the signs and symptoms (cf. Patel & Groen, 1986). This explanation constitutes a minimally sufficient explanation for the described signs and symptoms in a clinical case, and it is constructed with the help of domain experts (see Appendix 3.2). Canonical explanations are built upon high-level inferences derived from the signs and symptoms described in the text. The number of matching concepts (termed *Model concepts*) and links (termed *Model links*) are considered to be important indicators for the application of encapsulated knowledge in the pathophysiological explanations.

Inter-rater agreement for each of these procedures was higher than 90 %, and differences were resolved by discussions. The data were analysed using one-way analysis of variance (significance level was set at .05).

Results

Diagnostic accuracy

Figure 3.1 shows the mean accuracy of the diagnoses provided by the participants as a function of expertise level. Analysis of variance shows that the diagnostic accuracy of the cardiologists (the experts) was significantly higher than that of the subexperts [$F(1,18) = 46.10$, $MSE = 0.50$, $p < .0001$]. This result supports the general validity of the procedure and the selection of the participants. However, it is important to note that only two cardiologists provided a completely correct diagnosis on both cases.

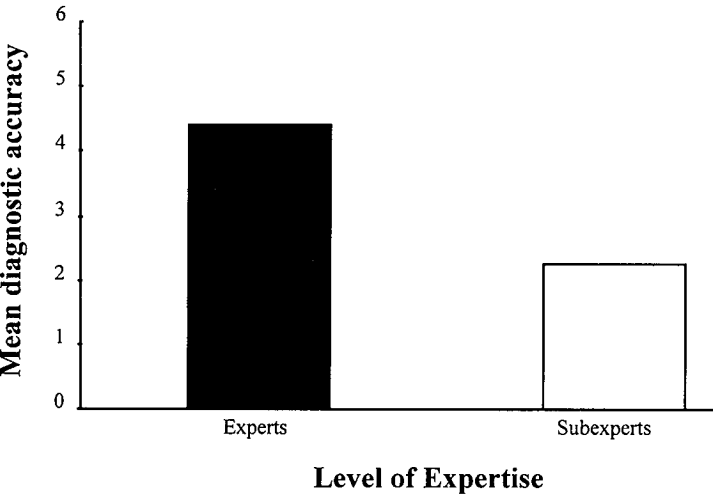


Figure 3.1 Mean diagnostic accuracy as a function of level of expertise.

Free Recall

The results for the mean number of propositions recalled and high-level inferences or Summaries are shown in Figure 3.2. Analysis of variance shows that the number of propositions recalled, as a function of level of expertise, was significant [$F(1,18) = 16.09$, $MSE = 23.25$, $p < .001$]. As was predicted, cardiologists reproduced fewer propositions from the clinical case description than did the neurologists.

Further, the presence of summaries in the recall protocols was looked at, because they are considered as important indicators whether or not encapsulated knowledge has been activated. Figure 3.2 shows the mean number of summaries within the participants' recall protocols. Overall differences between experts and subexperts were significant [$F(1,18) = 14.83$, $MSE = 0.92$, $p < .01$]. So, although cardiologists recalled fewer propositions from the text than did the subexperts, cardiologists used more encapsulating concepts in their recall.



Figure 3.2 Mean number of correctly recalled propositions and mean number of summaries by experts and subexperts.

Pathophysiological explanations

In order to analyse the pathophysiological explanations four measures were used. The mean number of concepts used to explain the signs and symptoms are displayed in Figure 3.3. The analysis of variance showed no significant difference between levels of expertise [$F(1,18) = 1.15$, $MSE = 11.13$, ns]. The mean number of links between the concepts is displayed in Figure 3.3. Again, no significant difference between levels of expertise was found [$F(1,18) = 0.46$, $MSE = 13.07$, ns]. The last two measures were based on the match of the explanations with the canonical models. Figure 3.3 shows the mean number of concepts that were identical or equivalent to those in the canonical explanations (the *model concepts*). The analysis of variance indicated a significant difference between levels of expertise [$F(1,18) = 9.02$, $MSE = 3.06$, $p < .01$]. Expert physicians introduced more model concepts in their explanations than did subexperts. Finally, the number of links that were identical with the canonical models were computed (the *model links*). The results for the mean number of identical links are presented in Figure 3.3. Again, there was a significant difference between levels of expertise [$F(1,18) = 8.46$, $MSE = 0.65$, $p < .01$]. Expert physicians also introduced more links that were identical with the canonical models than did subexperts.

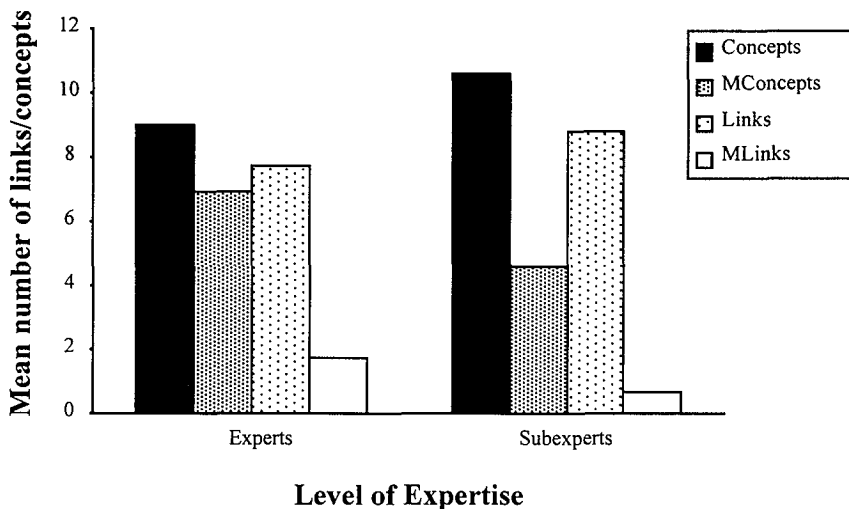


Figure 3.3 Mean number of concepts, links, model concepts (MConcepts), and model links (MLinks) in pathophysiological explanations by experts and subexperts.

Discussion

The pivotal role of encapsulating knowledge within the physician's domain of expertise has been repeatedly demonstrated in clinical case studies (e.g., Boshuizen, 1989; Coughlin & Patel, 1986; Patel & Medley-Mark, 1986; Schmidt & Boshuizen, 1993; Schmidt, Boshuizen, & Hobus, 1988). The experiment described in this article has been concerned with the role of encapsulated knowledge in the process of diagnosing clinical cases within and outside the physicians' domain of expertise. The employment of encapsulated knowledge was considered to be mainly confined to the physicians' domain of expertise and subexperts were therefore expected to fail in activating relevant encapsulating concepts. That is, subexperts have to resort to an elaborate biomedical processing approach in order to understand the patient's signs and symptoms. Expert physicians, on the other hand, were expected to rely on their encapsulated knowledge while diagnosing the clinical cases and therefore their recall and pathophysiological explanations would be more compendious than those of subexperts.

The results of the present study demonstrated that cardiologists (the experts) were indeed more accurate in their diagnoses than the neurologists (the subexperts). A mean diagnostic accuracy of 4.4 for the experts as compared to 2.25 for the subexperts clearly shows that diagnostic accuracy is a function of Expertise Level. However, the experts' diagnostic accuracy (about 73 % of the maximum score) also indicates that the expert group did not uncover all essential components of the correct diagnoses. This, and the fact that only two cardiologists provided a completely correct diagnosis on both cases, shows that the clinical cases employed were complex.

As has been outlined in the introduction section of this article, the measure of free recall plays an important role to support the theory of knowledge encapsulation (Boshuizen, 1989; Coughlin & Patel, 1986; Patel & Medley-Mark, 1986; Schmidt & Boshuizen, 1993; Schmidt, Boshuizen, & Hobus, 1988). Recall data of the present study showed differences between both groups that are in line with the predictions. Subexperts remembered more propositions from the clinical case description than did the experts, while experts applied more summaries or encapsulating concepts in their recall protocols than did subexperts. This finding supports the theory of knowledge encapsulation in that domain expertise is required in order to construct a clinical case representation based upon a small number of encapsulating concepts. Subexperts, on the other hand, were not able to process the clinical cases in an encapsulated mode and they had to resort to a deep (biomedical) processing approach in order to build a case representation.

Although the experts' performance on the measure of free recall was in line with our predictions, the pathophysiological data cannot be explained as straightforward. Subexperts were expected to relapse into long and detailed biomedical explanations for the described signs and symptoms. Expert physicians, on the other hand, were expected to provide short, and highly encapsulated explanations. The mean number of concepts and links used in their pathophysiological explanations, however, showed no significant differences between experts and subexperts. A fairly forthright explanation for this finding might be that expert physicians, as a result of the complexity of the cases (as was indicated by the measure of diagnostic accuracy), were not able to explain all signs and symptoms in an encapsulated mode. However, this explanation will not come up to the mark. If we take a closer look at the mean number of concepts and links used by the expert physicians (see Figure 3.3), we have to conclude that this is not in line with an *elaborate* biomedical explanation. Moreover, the overlap with the canonical explanations (i.e., the mean number of model concepts and links) was significantly higher for the experts than the subexperts, as was predicted by encapsulation theory. The percentage of encapsulating concepts in the cardiologists' pathophysiological explanations was very high (about 80 %). Neurologists' used substantially less encapsulating concepts in their explanations (about 45 %). This strongly suggests that the subexperts did relapse into a biomedical processing approach and, therefore, that the use of encapsulated knowledge is largely confined to the physicians' domain of expertise. Thus, although the pathophysiological data did not show differences in the quantity of concepts and links used, there were significant qualitative differences. However, this leaves us with an important question unanswered. If subexperts relapsed into an elaborate biomedical processing approach, why did not the present study reveal significant differences between experts and subexperts in the mean number of concepts and links? In order to answer this question, it is important to note that it is only the subexperts' performance that needs to be explained: Expert physicians' performance on all pathophysiological measures are completely in line with the predictions. A possible explanation might be that subexperts, as a consequence of the complexity of the cases, faced many loose ends while they processed the cases and were therefore not able to explain many of the described signs and symptoms; not in an encapsulated nor in a biomedical mode. In other words, subexperts, or for that matter any

expert, cannot explain things that are not understood. However, those things that were understood by the subexperts, were mostly explained by using biomedical concepts.

Interestingly, previous clinical case studies comparing medical experts with subexperts found, that, although experts outperformed subexperts on the measures of processing speed and diagnostic accuracy - the measure of free recall, however, failed to unearth differences between these groups (e.g., Joseph & Patel, 1990; Patel, Arocha, & Groen, 1986, 1990). This finding with regard to the measure of free recall, which is clearly at variance with the present study, is explained by Patel and colleagues by assuming that experts and subexperts alike have developed some kind of *generic expertise* for the relevant information in a clinical case. Generic expertise is concerned with the acquisition of adequate case representations, implying the existence of some schema or macrostructure representation (Kintsch, 1988; Van Dijk & Kintsch, 1983). This representation is based on the physicians' experience in practical problems, and guarantees the retention of the crucial facts of a clinical case within and outside their domain of expertise. Subexperts are therefore still able to recall the essential clinical case information. Expert physicians, on the other hand, have also acquired *specific expertise*, which is essential to achieve a high diagnostic accuracy. They possess the necessary domain knowledge that enables them to link non-salient cues or loose ends to the main diagnosis (Patel & Groen, 1991a). However, although the distinction between generic and specific expertise is very plausible, it cannot account for the findings of the present study, unless one assumes that our clinical cases demanded high levels of specific expertise in order to be recalled, which is not very likely. It is far more likely that the solution for the conflicting outcomes lied in the different ways the recall protocols were analysed. In contrast to the present study, Patel and colleagues classified, on the basis of experts' judgements, the recalled propositions as either disease relevant or disease irrelevant. The absence of significant differences in recall between experts and subexperts in their study, however, was only based on disease relevant propositions, excluding the other category. It is, therefore, not unlikely that subexperts recalled more irrelevant propositions than did expert physicians (cf. Groen & Patel, 1988; Patel, et al., 1991b).

It is important to note, that based on our data and those of previous studies some scepticism about the role of encapsulated knowledge as the sole cause of the relapse of expert physicians seems to be justified (e.g., Rikers, Boshuizen, & Schmidt, 1997; Rikers, Schmidt, & Boshuizen, in press). First of all, there are grounds for doubting the generality of the findings by Schmidt and Boshuizen (e.g., 1992, 1993) that suggest that extensive use of biomedical knowledge is a characteristic of non-expert reasoning. For instance, other studies have reported contrary results and emphasised the central role of biomedical knowledge in experts' reasoning. The main studies, which have indicated an increased role of biomedical knowledge in expert diagnosis, are those of Lesgold and colleagues on expertise in the domain of radiology (Lesgold, 1984; Lesgold, Rubinson, Feltovich, Glaser, Klopfer, & Wang, 1988). In these studies, it was found that correct interpretation of chest X-rays involved explicit use of anatomical and pathophysiological knowledge and increased use of such biomedical knowledge was especially evident in the experts' protocols.

However, it is possible that these findings are characteristic for the domain of radiology: Biomedical reasoning is here required to distinguish between alternative diagnoses that have similar surface manifestations (Gilhooly, 1996). Furthermore, no patient background information or contextual information was present in these studies and so short-cutting, based on encapsulating concepts, was less possible (Gilhooly, 1996). Therefore, it is not unlikely that since different studies have used tasks from different medical domains, certain domains permit rapid routes to diagnosis using clinical information, while others, such as radiology, may typically require extended biomedical reasoning even from experienced clinicians (cf. Gilhooly, McGeorge, Hunter, Rawles, Kirby, Green, & Wynn, 1997).

A second ground for doubting generality of the data by Schmidt and Boshuizen finds its source in the number of encapsulating concepts or summaries in the recall data. For instance, Schmidt & Boshuizen (1993) found that internists' (the experts) recall contained about 10 % encapsulations. The pattern that emerges from the present study is not really different. The proportion encapsulating concepts in recall is rather low: Cardiologists introduced on average about 17 % encapsulating concepts in their recall protocols. A fairly simple calculation tells us that 83 % of their recall was not in an encapsulated mode. Taken these proportions into consideration, it seems an overstatement to conclude that encapsulated knowledge is the only determinant of difference between experts and subexperts in the number of propositions recalled (cf. Van de Wiel, Schmidt, & Boshuizen, 1998). This issue is especially pertinent, because the intermediate effect, as it was originally found in clinical case studies, was solely contributed to the role of encapsulating concepts in experts' recall. The question that has to be answered, then, is how the findings on the measure of free recall of the present study must be interpreted. A first explanation might be that, as a consequence of the fact that expert physicians are not normally presented with problems outside their specialty, results were influenced by motivational factors. That is, specialists are not as inclined to write down everything they remember about a clinical case outside their speciality (Patel et al., 1990; Rikers, et al., *in press*). Secondly, as has been outlined before, subexperts might have had more difficulty to distinguish between case relevant and irrelevant information than experts might (Patel, et al, 1991b; Shanteau, 1992; Vincente & Wang, 1998). Subexperts processed, and hence recalled, much more case irrelevant information than did experts. Therefore, if recall is scored in terms of relevant information, the difference between experts and subexperts disappears (Groen & Patel, 1988; Patel, et al. 1986, 1990).

In this study, we have presented findings that shed some light on the role of encapsulated knowledge within and outside the expert physician's domain of expertise. In sum, this study largely supported the predictions made based on the theory of knowledge encapsulation. Still, it is not likely that encapsulating knowledge is the only constituent that can explain the differences between experts and subexperts as found in the present study. Therefore, it is necessary to investigate the contribution of encapsulated knowledge, and other constituents of the relapse of expert physicians outside their specialty.

References

- Blessing, S. B., & Anderson, J. R. (1996). How do people learn to skip steps? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22 (3), 576-598.
- Bordage, G. (1994). Elaborated knowledge: A key to successful diagnostic thinking. *Academic Medicine*, 69(11), 883-885.
- Boshuizen, H. P. A. (1989). *De ontwikkeling van medische expertise; een cognitief-psychologische benadering*. [On the development of medical expertise; a cognitive psychological approach]. Doctoral dissertation. Haarlem: Thesis.
- Boshuizen, H. P. A., & Schmidt, H.G. (1992). On the role of biomedical knowledge in clinical reasoning by experts, intermediates and novices. *Cognitive Science*, 16, 153-184.
- Charness, N. (1979). Components of skill in bridge. *Canadian Journal of Psychology*, 33, 1-16.
- Coughlin, L. D., & Patel, V. L. (1986, April). *Text comprehension and expertise in the domain of medicine*. Paper presented at the Meeting of the American Educational Research Association, San Francisco, CA.
- De Groot, A.D. (1946). *Het denken van den schaker* [Thinking processes in chess players]. Den Haag, The Netherlands: Noord Holland.
- De Groot, A. D. (1965). *Thought and choice in chess*. The Hague, The Netherlands: Mouton.
- Ericsson, K. A. (1985). Memory skill. *Canadian Journal of Psychology*, 39(2), 188-231.
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102(2), 211-245.
- Ericsson, K. A., Krampe, R. Th., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 3, 363-406.
- Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, 47, 273-305.
- Gilhooly, K.J. (1996). *Thinking: Directed, undirected and creative* (3rd ed.). London, UK: Academic Press.
- Gilhooly, K. J., McGeorge, P., Hunter, J., Rawles, J. M., Kirby, I. K., Green, C., & Wynn, V. (1997). Biomedical knowledge in diagnostic thinking: The case of electrocardiogram (ECG) interpretation. *European Journal of Cognitive Psychology*, 9(2), 199-223.
- Gilhooly, K.J. & Simpson, S. (1992). Deep knowledge in human medical expertise. In E. Keravnou (Ed.), *Deep models for medical knowledge in engineering*, pp 273-285. Amsterdam: Elsevier.
- Groen, G. J., & Patel, V. L. (1988). The relationship between comprehension and reasoning in medical expertise. In M. T. H. Chi, R. Glaser, & M. Farr (Eds.), *The nature of expertise*, (pp. 287-310). Hillsdale, NJ: Erlbaum.
- Jensen, A.R. (1990). Speed of information processing in a calculating prodigy. *Intelligence*, 14, 259-274.
- Joseph, G. M., & Patel, V. L. (1990). Domain knowledge and hypothesis generation in diagnostic reasoning. *Journal of Medical Decision Making*, 10, 31-46.
- Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, NJ: Erlbaum.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95(2), 163-182.
- Koedinger, K. R., & Anderson, J. R. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science*, 14, 511-550.
- Lesgold, A.M. (1984). Acquiring expertise. In J.R. Anderson & S.M. Kosslyn (Eds.), *Tutorials in learning and memory: Essays in honor of Gordon Bower*, (pp 31-60). San Francisco, CA: Freeman.

- Lesgold, A.M., Robinson, H., Feltovich, P.J., Glaser, R., Klopfer, D. & Wang, Y. (1988). Expertise in a complex skill: Diagnosing X-ray pictures. In M. T. H. Chi, R. Glaser, & M. Farr (Eds.), *The nature of expertise*, (pp. 311-342). Hillsdale, NJ: Erlbaum.
- McKeithen, K. B., Reitman, J. S., Rueter, H. H., & Hirtle, S. C. (1981). Knowledge organization and skill differences in computer programmers. *Cognitive Psychology*, 13, 307-325.
- Patel, V. L., Arocha, J. F., & Groen, G. J. (1986). Strategy selection and degree of expertise in medical reasoning. In C. Clifton (Ed.), *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*, (pp. 780-791). Hillsdale, NJ: Erlbaum.
- Patel, V. L., Evans, D. A., & Groen, G. J. (1989). Biomedical knowledge and clinical reasoning. In D. A. Evans & V. L. Patel (Eds.), *Cognitive science in medicine: Biomedical modeling*, (pp. 53-112). Cambridge, MA: MIT Press.
- Patel, V. L., & Groen, G. J. (1986). Knowledge based solution strategies in medical reasoning. *Cognitive Science*, 10, 91-116.
- Patel, V. L., & Groen, G. J. (1991a). The general and specific nature of medical expertise: A critical look. In A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits*, (pp. 93-125). New York: Cambridge University Press.
- Patel, V. L., & Groen, G. J. (1991b). Developmental accounts of the transition from medical student to doctor: Some problems and suggestions. *Medical Education*, 25(6), 527-535.
- Patel, V. L., Groen, G. J., & Arocha, J. F. (1990). Medical expertise as a function of task difficulty. *Memory and Cognition*, 18(4), 394-406.
- Patel, V. L., & Kaufman, D. R. (1995). Clinical reasoning and biomedical knowledge: Implications for teaching. In J. Higgs & M. Jones (Eds.), *Clinical reasoning in the health professions*, (pp. 117-128). Oxford, UK: Butterworth Heinemann.
- Patel, V.L., & Medley-Mark, V. (1986). Relationship between representation of textual information and underlying problem representation in medicine. *Cognitive research series*. Montreal: McGill University, Centre for Medical Education. (CME86-CS1).
- Reitman, J. S. (1976). Skilled perception in Go: Deducing memory structures from inter-response times. *Cognitive Psychology*, 8, 336-356.
- Rikers, R. M. J. P., Boshuizen, H. P. A., & Schmidt, H. G. (1997, March). *Nonroutine problem solving by medical experts*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, ILL.
- Rikers, R. M. J. P., Schmidt, H. G., & Boshuizen, H. P. A. (in press). Knowledge encapsulation and the intermediate effect. *Contemporary Educational Psychology*.
- Schmidt, H. G., & Boshuizen, H. P. A. (1992). Encapsulation of biomedical knowledge. In D. A. Evans & V. L. Patel (Eds.), *Advanced models of cognition for medical training and practice*. New York, NY: Springer Verlag.
- Schmidt, H. G., & Boshuizen, H. P. A. (1993). On the origin of intermediate effects in clinical case recall. *Memory and Cognition*, 21(3), 338-351.
- Schmidt, H. G., Boshuizen, H. P. A., & Hobus, P. P. M. (1988). Transitory stages in the development of medical expertise: The "intermediate effect" in clinical case representation studies. In: *Proceedings of the Cognitive Science Society Meeting*, Lawrence Erlbaum, Hillsdale, NJ.
- Schmidt, H. G., Boshuizen, H. P. A., & Norman, G. R. (1992). Reflections on the nature of expertise in medicine. In E. Keravnou (Ed.), *Deep models for medical knowledge engineering*. Amsterdam, The Netherlands: Elsevier.
- Shanteau, J. (1992). How much information does an expert use? Is it relevant? *Acta Psychologica*, 81(1), 75-86.

- Simpson, S. A., & Gilhooly, K. J. (1997). Diagnostic thinking processes: Evidence from a constructive interaction study of electrocardiogram (ECG) interpretation. *Applied Cognitive Psychology*, 11(6), 543-554.
- Spillich, G.J., Vesonder, G.T., Chiesi, H.L., & Voss, J.F. (1979). Text processing of domain-related information for individuals with high and low domain knowledge. *Journal of Verbal Learning & Verbal Behavior*, 18, 275-290.
- Van de Wiel, M.W.J., Schmidt, H.G., & Boshuizen, H.P.A. (1998). A failure to reproduce the intermediate effect in clinical case recall. *Academic Medicine*, 73, 894-900.
- Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Vicente, K. J., & Wang, J. H. (1998). An ecological theory of expertise effects in memory recall. *Psychological Review*, 105(1), 33-57.

Appendix 3.1

Case of Bacterial Endocarditis

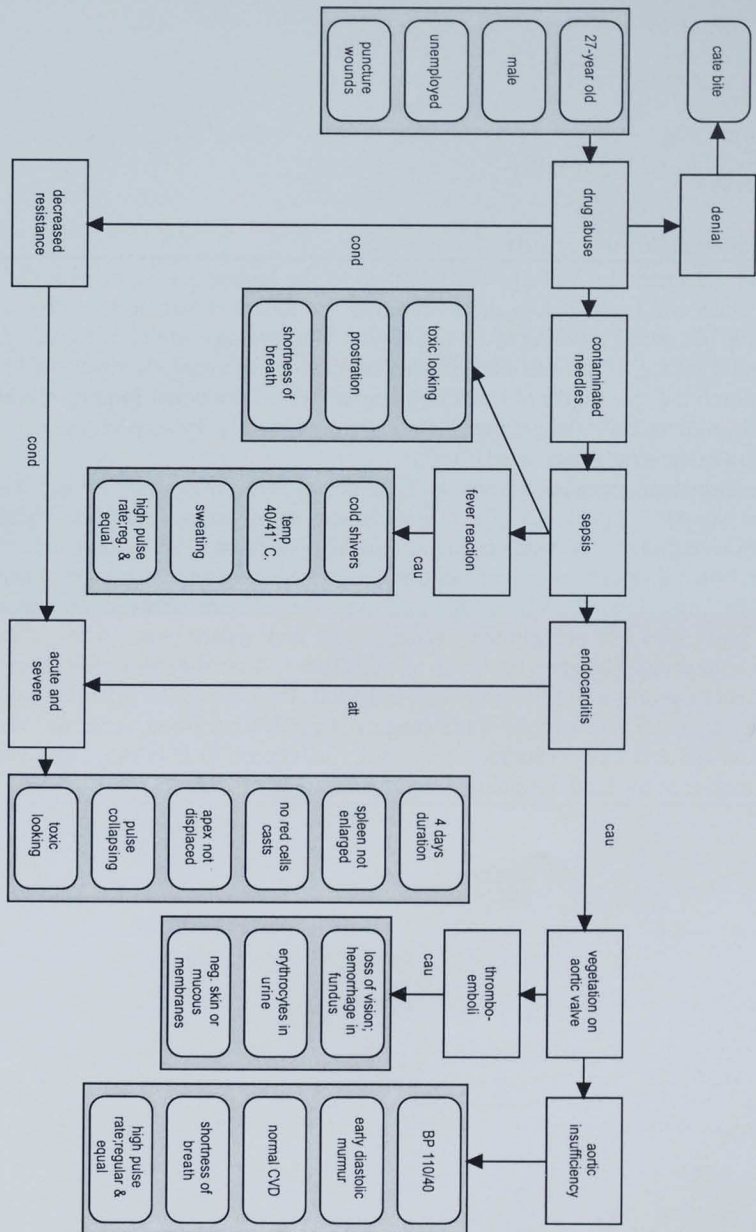
This 27 year old unemployed male was admitted to the emergency room with the complaint of shaking chills and fever of four days' duration. He took his own temperature and it was recorded at 40° C. on the morning of his admission. The fever and chills were accompanied by sweating and a feeling of prostration. He also complained of some shortness of breath when he tried to climb the two flights of stairs in his apartment. Functional inquiry revealed a transient loss of vision in his right eye which lasted approximately 45 seconds on the day before his admission to the emergency ward.

Physical examination revealed a toxic looking young man who was having a rigor. His temperature was 41° C., pulse 120, BP 110/40. Mucus membranes were pink. Examination of his limbs showed puncture wounds in his left antecubital fossa. The patient volunteered that he had been bitten by a cat at a friend's house about a week before admission. There were no other skin findings. Examination of the cardiovascular system showed no jugular venous distention, pulse was 120 per minute, regular, equal and synchronous. The pulse was also noted to be collapsing. The apex beat was not displaced. Auscultation of his heart revealed a 2/6 early diastolic murmur in the aortic area and funduscopy revealed a flame shaped hemorrhage in the left eye. There was no splenomegaly. Urinalysis showed numerous red cells but there were no red cell casts. (Source: Patel, V.L., & Groen, G.J. (1986). Knowledge-based solution strategies in medical reasoning. *Cognitive Science*, 10, 91-116)

Appendix 3.2

Canonical explanation for the Endocarditis case

A canonical explanation of a case of bacterial endocarditis. Frames with rounded corners represent signs and symptoms present in the case, and frames with rectangular corners represent explanatory concepts. The links between concepts have labels such as, *cau.* (causal), *att.* (attribute), *loc.* (location), *spec.* (specification), *temp.* (temporal), *cond.* (conditional), *iden.* (identity).



Chapter 4

Knowledge Encapsulation and the Intermediate Effect¹

The present study explored the role of so-called encapsulated knowledge in diagnosing clinical cases outside the expert physicians' domain of expertise. Neurologists, second-year sixth-year and medical students were required to diagnose, recall, and explain the signs and symptoms of two cardiological and two pulmonological clinical case descriptions. Our experiment showed that neurologists diagnosed these clinical cases faster and more accurately than second-year and sixth-year medical students. An inverted U-shaped relationship with levels of expertise was found in recall and pathophysiological protocols: Sixth-year medical students remembered more information from the cases and produced more elaborated explanations for the described signs and symptoms than both other groups. The proportion of encapsulating concepts in recall and pathophysiological explanations, on the other hand, increased with levels of expertise. This pattern is similar to that found in previous studies on clinical case representations using only cases within the expert physicians' domain of expertise. Therefore, these results suggest, that expert physicians process clinical case descriptions both within and outside their domain of expertise in essentially the same way.

¹ Rikers, R.M.J.P., Schmidt, H.G., Boshuizen, H.P.A. (1998, April). *Intermediate effect in recall of nonroutine clinical cases*. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA. This chapter will appear as Rikers, R.M.J.P., Schmidt, H.G., Boshuizen, H.P.A. (in press). Knowledge Encapsulation and the intermediate effect, *Contemporary Educational Psychology*.

Introduction

Chess Masters can recall almost perfectly a chess position that they have seen for just a few seconds (Chase & Simon, 1973; De Groot, 1965; Gobet & Simon, 1996a, 1996b, 1996c). Waiters and waitresses show a superior memory performance for drink and dinner orders (Bennett, 1983; Ericsson & Polson, 1988a, 1988b). Mental calculators can rapidly multiply large numbers in their head and also have exceptional memory for digits (Ericsson, 1985; Jensen, 1990). But also research in other domains such as bridge (Charness, 1979), go (Reitman, 1976) and computer programming (McKeithen, Reitman, Rueter, & Hirtle, 1981) generally demonstrate a superior performance of domain experts. According to Vincente and Wang (1998) there are at least 51 studies in at least 19 different domains demonstrating the superior memory performance of experts on meaningful stimuli.

However, in the domain of medicine findings have been reported that seem to imply that the relationship between expertise level and recall is not as straightforward as in the aforementioned domains. A consistent finding in clinical case representation studies is the so-called *intermediate effect* (e.g., Boshuizen, 1989; Claessen & Boshuizen, 1985; Patel & Groen, 1986; Patel & Medley-Mark, 1986; Schmidt & Boshuizen, 1993; Schmidt, Boshuizen, & Hobus, 1988). In these studies, participants of different levels of medical expertise were instructed to read a text describing a patient's history, complaint(s) and some additional findings such as physical examination and laboratory data (see Appendix 4.1). After the text was removed, each subject was asked to provide a diagnosis and to write down as much of the text as they could remember. Finally, the subjects were asked to explain in writing the signs and symptoms displayed in the clinical case (the so-called pathophysiological explanation). The outcome typically was, that medical students of an *intermediate level* of expertise (i.e., advanced students), remembered more, and explained the signs and symptoms of a clinical case in a more elaborated fashion than either expert physicians or novices. So, rather than a monotonically increasing function with increasing levels of expertise, an inverted U-shaped relation was found. This counterintuitive phenomenon of the recall and pathophysiological data is difficult to explain in the light of the aforementioned domains, but also contemporary theories of text processing. These theories generally assume that prior knowledge facilitates the process of encoding and retrieval of textual information (e.g., Ericsson & Kintsch, 1995; Graesser & Clark, 1985; Kintsch, 1988; Voss & Bisanz, 1985).

Schmidt and Boshuizen (1993) have demonstrated that the occurrence of an intermediate effect is dependent on the amount of time available for processing a clinical case description. In one of their studies, participants of different levels of expertise were subdivided into three groups who studied a clinical case description under different time constraints (3.30 min., 1.15 min., or 30 sec.). The results showed an increase in diagnostic accuracy with levels of expertise for all study times, but this relationship was not found for the recall data. At the longest and the medium study times, recall data showed a clear intermediate effect (i.e., advanced medical

students remembered more of the description than novices and expert physicians). The shortest study time, on the other hand, displayed a clear expert superiority effect in the recall.

Schmidt and Boshuizen (1992, 1993) have provided a possible explanation for the intermediate effect. They suggest that novices, intermediates and expert physicians apply qualitatively different forms of medical knowledge to deal with their task. According to Schmidt and Boshuizen the novices' medical knowledge is best characterized by little knowledge of biomedicine, and even less about the manifestations of a disease in a patient. Intermediates, however, have acquired extensive biomedical knowledge from textbooks and lectures, but as the novices they have not got extensive clinical knowledge.² Therefore, if advanced students have to diagnose a clinical case, they will employ large amounts of biomedical knowledge to account for the described signs and symptoms. They will consciously relate the signs and symptoms they encounter to concepts in their pathophysiological knowledge base to reason through the case, activating whatever knowledge is available. The expert physicians' knowledge base, however, is different from that of novices and intermediates in that biomedical knowledge only plays a minor and implicit role in normal diagnostic reasoning (Boshuizen & Schmidt, 1992; Gilhooly, 1996; Gilhooly & Simpson, 1992; Simpson & Gilhooly, 1997). As a result of extensive practice and the confrontation with actual patients, the physicians' biomedical knowledge has become linked with, or *encapsulated* under, a limited number of clinically relevant concepts, that have the same explanatory power as the elaborate biomedical structure. This restructuring, which takes place in the course of the development from novice to expert physician, leads eventually to abbreviations in lines of reasoning (cf. Koedinger & Anderson, 1990; Elio & Scharf, 1990). It is important to note, that the activation or retrieval of encapsulating concepts is a dynamic process, taking place while the clinical case information is processed. In other words, information about signs and symptoms in a case, point towards, and therefore activate, specific "prestored" encapsulating concepts.

For instance, suppose that one would require a cardiologist and an advanced student to study a clinical case description that contains, among other things, the following information: *A 43-year old man with a rapid and weak apex beat; pulse is irregular and weak; tenderness over precordium, percussion is negative, auscultation reveals first sound of heart resembling second heart sound, being high pitched and wanting in muscular quality* (Thomas, 1989). A cardiologist will tend to recall this information as "Myocarditis" or "Inflammation of the myocardium." This is an (high-level) inference from the text that summarizes or encapsulates the provided information. An advanced student, on the other hand, will likely reproduce the information as it is written, without high-level inferences as made by a cardiologist.

² Clinical knowledge concerns the ways in which a disease can manifest itself in patients. Biomedical knowledge, or basic science knowledge, pertains to the (pathological) processes underlying the manifestations of disease. It incorporates knowledge about domains such as biochemistry, microbiology, and physiology (Patel, Evans, & Groen, 1989; Patel & Kaufman, 1995).

Therefore, the recall performance of advanced students is not really better than that of an expert physician, unless we want to equate more with better (Patel & Groen, 1991b).

Note that knowledge encapsulation explains the above-mentioned disappearance of the intermediate effect if processing time of a clinical case is severely reduced. Under this condition advanced students fail to construct an adequate clinical case representation, because their elaborate approach takes considerably more time than the physician's approach of constructing a representation through the application of a small number of encapsulating concepts. Students do not have sufficient time to engage in deeper pathophysiological reasoning and so lost their advantage in terms of text recall (for a more detailed discussion, see Boshuizen & Schmidt, 1992; Gilhooly, 1996; Patel & Groen, 1991b; Schmidt & Boshuizen, 1993; Schmidt, Boshuizen, & Norman, 1992).

Evidence for knowledge encapsulation is not only found in recall data, but also in post-hoc pathophysiological explanations (Boshuizen & Schmidt, 1992; Schmidt & Boshuizen, 1992, 1993). Expert physicians who were asked to explain the signs and symptoms in a clinical case description, referred primarily to clinically relevant concepts. Advanced students, on the other hand, used primarily biomedical concepts to account for the signs and symptoms. Interestingly, Patel, Arocha, and Groen (1986, 1990) found that subexperts (i.e., expert physicians diagnosing clinical cases outside their specialty), also applied more biomedical concepts as compared to experts diagnosing problems within their domain of expertise. Schmidt and Boshuizen assume that biomedical knowledge, although not used by experts in routine cases, remains easily accessible when the need arises. They have postulated that the different knowledge structures acquired in the course of the development towards expertise are stored into long-term memory, and are still accessible when more recently acquired, encapsulated structures, fail to produce an adequate representation of a clinical problem (Boshuizen & Schmidt, 1992; Schmidt & Boshuizen, 1993; Schmidt, Boshuizen, & Norman, 1992).

In the present study, Neurologists (subexperts in this study), second-year and sixth-year medical students were required to diagnose, recall and explain the signs and symptoms of four clinical cases. Two clinical cases were within the domain of pulmonology, the other two were cardiological cases. Based on the assumptions that the development of expertise in a domain is characterized by encapsulation of biomedical knowledge into clinically relevant concepts, and that the application of these concepts is confined to the domain of expertise, it is predicted that expert physicians' performance will be comparable to, and not significantly different from, that of students of an intermediate level of expertise. In other words, it is assumed that if expert physicians are confronted with cases outside their domain of expertise they will relapse into the deep or elaborate biomedical processing approach which is typical for advanced medical students.

Method

Participants

The participants were 16 medical students (8 second-year and 8 sixth-year) of Maastricht University, and 8 physicians from three hospitals in the southeastern part of the Netherlands. The second-year medical students (the novice group) had no, or very little, experience in hospitals and their knowledge of disease was therefore principally from textbooks and lectures. The sixth-year medical students (the intermediate group) had completed an in-hospital training under supervision of senior residents and physicians. They had almost completed all of the requirements for obtaining their medical degrees as physicians. The neurologists (the subexpert group) were practitioners with at least an MD degree and Board Certification in their specialty and with at least 4 years of experience ($M = 16$). All participants received a financial compensation for their participation.

Material

The materials consisted of four clinical case descriptions presented on a computer screen and two blank response sheets after each case. The clinical cases consisted of two cardiological and two pulmonary cases and were constructed by two domain experts. The cases were based on actual patients treated at the Maastricht University Hospital. Each clinical case description reported some contextual information, the complaint(s), findings from history taking and physical examination, relevant laboratory data and some additional findings (e.g., X-rays and ECG's). All cases (Aortic Valve Insufficiency, Congestive Heart Failure, Small Cell Lung Cancer, and Legionnaires' Disease) were outside the domain of expertise of the neurologists. The four case descriptions were about one page in length, and consisted of 94, 80, 104, and 105 propositions respectively. The translated text of the case of congestive heart failure is provided in Appendix 4.1.

Procedure

The participants were successively presented four written descriptions of clinical cases. They were instructed to read each case description carefully, and to provide a diagnosis (students and physicians alike are familiar with written cases, because they are often used in medical training and practice). Subsequently, they had to write down everything they could remember about a clinical case description. Finally, they had to explain the described signs and symptoms: The so-called pathophysiological explanation. The participants knew beforehand that a recall task and a pathophysiological explanation would follow each case. All participants were tested individually.

The cases were divided into a number of segments of related information (cf. Joseph & Patel, 1990; Lemieux & Bordage, 1992). The segments were presented successively on a portable Macintosh computer (Powerbook 180). New segments appeared on the screen when the mouse button was pressed. It was impossible to look again at a previous segment. Each participant had to evaluate all cases. Carry-over effects were controlled by randomizing the presentation order of the cases. The participants were free to spend as much time as they needed to complete the different assignments. The time spent reading a case was registered by the computer. To enable the participants to get used to their tasks, the experimental cases were preceded by one sample case.

Analysis

The accuracy and completeness of the provided diagnosis for each case was scored by two expert physicians on a scale ranging from 0 (completely incorrect diagnosis) to 6 (completely correct diagnosis). For example, the complete and accurate diagnosis of one of the pulmonary cases was: Small cell lung cancer with metastases in the mediastinal lymph nodes. If a participant produced a diagnosis containing the term "Lung cancer" (or an equivalent) 3 points were given, because it is the most essential part of the correct diagnosis. One point was given for each of the terms "Small Cell," and "Metastases." Finally, a point could be earned by providing the correct location of the metastases. If the correct diagnosis was part of a differential diagnosis, it was considered correct.

The protocol analysis of recall and pathophysiological data consisted of segmenting the protocols into propositions based on the work of Kintsch (1974, 1988). The essential element in propositional analysis is the segmentation of a clinical case into individual propositions corresponding to discrete idea units in the text (see for a more detailed discussion of propositional analysis, Patel & Groen, 1986). For each proposition in the free recall, it was decided whether or not it matched any proposition in the text. The identified propositions were divided into three types: literal, paraphrased, and inferred propositions, which have been transformed in some way. Evidence for encapsulation of signs and symptoms was explored by counting the number of high-level inferences in the recall protocols. High-level inferences were included in the analysis to the extent that they could be matched to a combination of propositions in the original text (Coughlin & Patel, 1986). Inferences referring to only one proposition in the text were not considered as encapsulations. To distinguish high-level inferences from inferences based on only one proposition, they will be further referred to as *summaries*. The total recall score was based on the number of literal matching propositions, paraphrased propositions, inferred propositions, and summaries.

From each of the pathophysiological explanations a conceptual network was constructed, that formalized the semantic relations underlying the disease. Four measures were used to score the pathophysiological protocols. First, the number of

different concepts used in the participants' explanations was counted. Second, the number of relations or links between these concepts were counted. The two other pathophysiological measures were based on the overlap with a so-called canonical explanation of the signs and symptoms (Patel & Groen, 1986). A canonical explanation constitutes a minimally sufficient explanation for all the signs and symptoms described in a clinical case and it is constructed with the help of domain experts. Canonical explanations are built upon high-level inferences derived from the signs and symptoms in the text (See Appendix 4.2). The number of matching concepts (termed *model concepts*) and links (termed *model links*) are therefore considered to be important indicators for the application of encapsulated knowledge in the pathophysiological explanations.

Interrater agreement for each of these procedures was higher than 95 %. The data were analyzed using one-way analysis of variance (significance level was set at .05). In addition, the Student-Newman-Keuls test was used to make pairwise comparisons between the different levels of expertise (significance level was set at .05).

Results

Case reading times

Table 4.1 shows the mean reading times in seconds for the different levels of expertise. Analysis of variance indicated that the time each participant spent reading a case was associated with levels of expertise [$F(2,21) = 4.80, p < .05$]. Newman-Keuls tests indicated that neurologists were faster than both second and sixth-year medical students. There were no significant differences between second-year and sixth-year students.

Table 4.1
Mean Reading Times in Seconds as a Function of Level of Expertise

Levels of expertise	<i>M</i>	<i>SD</i>
Second-year students	295.82	91.80
Sixth-year students	322.38	61.79
Neurologists	211.21	67.88

Diagnostic accuracy

The accuracy of the provided diagnoses is associated with levels of expertise [$F(2,21) = 28.28, p < .0001$]. The mean accuracy increased with expertise level (see Table 4.2). Newman-Keuls tests revealed that neurologists provided more accurate diagnoses than second-year students and sixth-year students. Pairwise comparisons between neurologists, second-year and sixth-year medical students indicated a posi-

tive linear relationship with levels of expertise. It is important to note that none of the participants provided a completely accurate diagnosis. These results suggest, therefore, that none of the participants perceived the cases as routine problems.

Table 4.2
Mean Accuracy of Diagnoses as a Function of Levels of Expertise

Levels of expertise	<i>M</i>	<i>SD</i>
Second-year students	0.78	0.37
Sixth-year students	2.34	1.01
Neurologists	3.28	0.47

Free Recall

The results for the mean number of correctly recalled propositions are presented in Table 4.3. Analysis of variance showed that the number of propositions recalled, as a function of levels of expertise, was significant [$F(2,21) = 3.99, p < .05$]. Newman-Keuls tests showed no significant difference between second-years students and neurologists, but the differences between these groups and sixth-year students were significant. The inverted U-shaped curve, indicating an intermediate effect, seems to provide the best fit for these data.

Table 4.3
Mean Number of Propositions Recalled as a Function of Levels of Expertise

Levels of expertise	<i>M</i>	<i>SD</i>
Second-year students	23.38	7.66
Sixth-year students	36.84	7.34
Neurologists	25.31	14.34

Further, the presence of summaries in the recall protocols was looked at, because they are considered as important indicators whether or not encapsulated knowledge had been applied. Table 4.4 shows the proportion of summaries within the participants' recall protocols. Overall differences between levels of expertise were significant [$F(2,21) = 17.30, p < .001$]. Newman-Keuls tests showed a significant difference between the neurologists and both student groups. So, although their total number of propositions recalled is less than sixth year students, neurologists introduced more encapsulating concepts in their recall than both second-year and sixth-year students.

Table 4.4
Mean Proportion of Summaries as a Function of Levels of Expertise

Levels of expertise	<i>M</i>	<i>SD</i>
Second-year students	2.15	1.99
Sixth-year students	3.66	1.62
Neurologists	7.53	2.02

Pathophysiological explanations

In order to analyze the pathophysiological explanations four measures were used. First, we looked at the total number of concepts used to explain the signs and symptoms. The mean number of concepts used to explain the signs and symptoms are displayed in table 4.5. The analysis of variance showed a significant difference between levels of expertise [$F(2,21) = 6.44, p < .01$].

Table 4.5
Mean Number of Concepts as a Function of Levels of Expertise

Levels of expertise	<i>M</i>	<i>SD</i>
Second-year students	7.19	3.80
Sixth-year students	14.50	6.02
Neurologists	7.34	3.80

Second, the number of relations or links between those concepts was analyzed. The mean number of links between the concepts is displayed in table 4.6. Again, a significant difference with levels of expertise was found [$F(2,21) = 5.58, p < .05$]. Newman-Keuls tests for the total number of concepts and links, showed that sixth-year students produced more elaborate (more links and concepts) protocols than neurologists and second-year students.

Table 4.6
Mean Number of Links as a Function of Levels of Expertise

Levels of expertise	<i>M</i>	<i>SD</i>
Second-year students	5.81	3.43
Sixth-year students	12.63	5.91
Neurologists	6.19	4.05

The last two measures were based on the match of the explanations with the so-called canonical models. Table 4.7 shows the proportion of concepts used that were identical or equivalent to those in the canonical explanation. The analysis of variance indicated a significant difference between levels of expertise [$F(2,21) = 9.87, p < .001$].

Table 4.7*Mean Proportion of Model Concepts as a Function of Levels of Expertise*

Levels of expertise	<i>M</i>	<i>SD</i>
Second-year students	32.66	10.41
Sixth-year students	41.53	6.20
Neurologists	56.49	14.36

Finally, the proportion of links that were identical with the canonical model was computed. The results for the proportion of identical links are presented in Table 4.8. There was a significant difference between levels of expertise [$F(2,21) = 11.02, p < .001$].

Newman-Keuls tests for the proportion of model concepts and model links showed that neurologists produced more overlap with the canonical explanations than second-year and sixth-year students.

Table 4.8*Mean Proportion of Model Links as a Function of Levels of Expertise*

Levels of expertise	<i>M</i>	<i>SD</i>
Second-year students	4.26	6.91
Sixth-year students	7.09	3.89
Neurologists	20.40	9.94

Discussion

The present study investigated the role of encapsulated knowledge in the process of diagnosing clinical cases outside the expert physician's domain of expertise. Previous studies (using cases within the expert physicians' domain of expertise) have shown that although physicians performed better than students on the measure of diagnostic accuracy - the total number of propositions recalled, and the elaborateness of the pathophysiological explanations showed a nonlinear relationship with levels of expertise (e.g., Boshuizen, 1989; Boshuizen & Schmidt, 1992; Schmidt & Boshuizen, 1993). Based on the theory of knowledge encapsulation a number of predications were made with regard to the performance of expert physicians and medical students. It was predicted that expert physicians cannot solve a problem outside their specialty in an encapsulated mode, and therefore they would perform at an intermediate level of expertise. This relapse of expert physicians should be expressed in a performance, on all measures, that is not significant different from that of sixth-year medical students, the intermediates in this study.

The results of the present study clearly demonstrated that expert physicians do not show a performance at an intermediate level of expertise when they are

confronted with clinical cases outside their domain of expertise. First, experts were faster and more accurate in evaluating the clinical cases outside their domain of expertise than second-year and sixth-year medical students. Second, the recall and pathophysiological data showed intermediate effects analogical to those found in clinical case studies using only cases within the physicians' domain of expertise (e.g., Schmidt & Boshuizen, 1992, 1993).

A fairly straightforward explanation for these findings would be that experts process cases within and outside their domain of expertise in essentially the same way. Qualitative differences (i.e., differences in the number of encapsulating concepts), as found between expert physicians and intermediates solving problems within the physicians' domain of expertise do also apply comparing these groups under more "non-routine conditions." Therefore, the presence of an intermediate effect, as it was originally found in recall and pathophysiological data, seems to be relatively insensitive to the nature of the cases used.

Evidence to support this explanation is also found in a recent study of Rikers, Schmidt, & Boshuizen (1997), who compared cardiologists and pulmonologists diagnosing each other's clinical cases. It was predicted, as in this study, that the expert physicians' recall should be more extensive outside their specialty, because they cannot rely on their encapsulated knowledge. Nonetheless, their study showed no significant differences in recall data between cardiologists and pulmonologists. In addition, Patel and colleagues also found no differences in recall of relevant and irrelevant propositions between expert physicians (cardiologists) and subexperts (surgeons and psychiatrists). Again, this seems to indicate that expert physician's performance is relatively insensitive for the nature of the cases used (Patel & Groen, 1986; Patel, Groen, & Arocha, 1990). Surely, the findings of the present study do not suggest that expert physicians perform outside their domain as well as inside, as is substantiated by the measure of diagnostic accuracy. The diagnostic performance of the neurologists had a mean accuracy of about 50 % of the maximum score. This clearly indicates that they did not uncover all crucial components of the correct diagnoses.

It is important to note that based on our data and those of previous studies, some skepticism about the role of encapsulated knowledge as the single cause of the intermediate effect seems to be justified. If we reconsider the recall data of this study, we have to conclude that the proportion of encapsulating concepts within the recall protocols is rather low - albeit significantly different from second-year and sixth-year medical students. Neurologists introduced less than 8 % summaries in their recall. So, more than 90% of their recall were literal, paraphrased, and (low-level) inferred propositions. If we take these proportions into consideration, we have to conclude that the recall protocols are not "characterized" by encapsulating concepts (cf. Van de Wiel, Boshuizen, & Schmidt, 1998). The pathophysiological explanations, however, seem to supply more support for the pivotal role of encapsulated knowledge. Neurologists introduced in their explanations about 56% model concepts and 20% model links, which is more than either second-year or sixth-year

medical students. On the other hand, more than 40% of the concepts and about 80% of the links in their explanations did not overlap with the canonical models. Again, but a little less than the recall data, we have to conclude that the neurologists' pathophysiological explanations are not "dominated" by encapsulated knowledge. Note that this pattern is similar to that found in previous clinical case studies. For example, Schmidt and Boshuizen (1993) found that expert physicians (internists) introduced about 10 % high-level inferences in their recall, and about 30 % model concepts in their pathophysiological explanations. Therefore, the pathophysiological data, but especially the recall data, seem to indicate that an explanation of the intermediate effect solely based on knowledge encapsulation cannot be the entire story (cf. Patel & Groen, 1991b).

The question, then, is how the intermediate effect must be interpreted, if knowledge encapsulation is not the only cause? Possibly, the inverted U-shaped curve found in the recall and pathophysiological data may also reflect a difference in the way the task is perceived. That is, expert physicians are not as inclined to write down everything (especially case information that is of no significance) they remember of a clinical case description than students do. Evidence to support this hypothesis, however, is primarily anecdotal. Phrases like, "Must I really write down *everything* I remember"? or "Should I explain *all* signs and symptoms"?, were sometimes stated by expert physicians but never by medical students.

Another plausible explanation is provided by Patel and her colleagues (Patel & Groen, 1991a, 1991b; Patel, Groen, & Arocha, 1990). They assume that both expert physicians and subexperts have developed some kind of generic expertise for relevant information in a clinical case. Generic expertise is concerned with the acquisition of adequate case representations and it implies the existence of some schema or macrostructure representation (Van Dijk & Kintsch, 1983). This representation is based on the physicians' experience in practical problems, and guarantees the retention of crucial facts in a clinical case within or outside their domain of expertise. Therefore, subexperts are still able to separate the wheat from the chaff, and only recall the essential information of a clinical case. Domain experts, on the other hand, have also acquired specific expertise that is essential to achieve a high diagnostic accuracy. They possess the necessary domain knowledge that enables them to link the nonsalient cues or loose ends to the main diagnosis (Patel & Groen, 1991a). However, advanced medical students have more difficulty to distinguish between relevant and irrelevant information than experts and subexperts do (Shanteau, 1992; Vincente & Wang, 1998). Therefore, if recall is scored in terms of relevant and irrelevant information, the intermediate peak disappears: Students process too much irrelevant information, whereas experts and subexperts do not (Patel, 1984; Patel & Groen, 1991b). This seems to indicate that the criterion for expert performance is superior diagnostic accuracy, not the reproduction of irrelevant details of a clinical case (Ericsson & Lehmann, 1996; Groen & Patel, 1988).

In sum, although this study did not support the hypothesis of encapsulation theory that subexperts relapse into an intermediate level of expertise, it did show that

the intermediate effect, as it was originally found in recall and pathophysiological data, seems to be independent of the nature of the cases used (i.e., within or outside the expert physician's domain of expertise). Further research is necessary in order to unearth the contribution of encapsulated knowledge, and other constituents of the intermediate effect within and outside the physician's domain of expertise.

References

- Bennett, H. L. (1983). Remembering drink orders: the memory skill of cocktail waitresses. *Human Learning*, 2, 157-169.
- Boshuizen, H. P. A. (1989). *De ontwikkeling van medische expertise; een cognitief-psychologische benadering*. [On the development of medical expertise; a cognitive psychological approach]. Doctoral dissertation. Haarlem: Thesis.
- Boshuizen, H. P. A., & Schmidt, H. G. (1992). On the role of biomedical knowledge in clinical reasoning by experts, intermediates and novices. *Cognitive Science*, 16, 153-184.
- Charness, N. (1979). Components of skill in bridge. *Canadian Journal of Psychology*, 33, 1-16.
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55-81.
- Claessen, H. F. A., & Boshuizen, H. P. A. (1985). Recall of medical information by students and doctors. *Medical Education*, 19, 61-67.
- Coughlin, L. D., & Patel, V. L. (1986, April). *Text comprehension and expertise in the domain of medicine*. Paper presented at the Meeting of the American Educational Research Association, San Francisco, CA.
- De Groot, A. D. (1965). *Thought and choice in chess*. The Hague, The Netherlands: Mouton.
- Elio, R., & Scharf, P. B. (1990). Modeling novice-to-expert shifts in problems-solving strategy and knowledge organization. *Cognitive Science*, 14, 579-639.
- Ericsson, K. A. (1985). Memory skill. *Canadian Journal of Psychology*, 39(2), 188-231.
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102(2), 211-245.
- Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, 47, 273-305.
- Ericsson, K. A., & Polson, P. G. (1988a). An experimental analysis of the mechanisms of a memory skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(2), 305-316.
- Ericsson, K. A., & Polson, P. G. (1988b). Memory for restaurant orders. In M. T. H. Chi, R. Glaser, & M. Farr (Eds.), *The nature of expertise* (pp. 23-70). Hillsdale, NJ: Erlbaum.
- Gilhooly, K.J. (1996). *Thinking: Directed, undirected and creative* (3rd ed.). London, UK: Academic Press.
- Gilhooly, K.J. & Simpson, S. (1992). Deep knowledge in human medical expertise. In E. Keravnou (Ed.), *Deep models for medical knowledge in engineering*, pp 273-285. Amsterdam: Elsevier.
- Gobet, F., & Simon, H. A. (1996a). Recall of random and distorted chess positions: Implications for the theory of expertise. *Memory and Cognition*, 24(4), 493-503.
- Gobet, F., & Simon, H. A. (1996b). Recall of rapidly presented random chess positions is a function of skill. *Psychonomic Bulletin and Review*, 3(2), 159-163.

- Gobet, F., & Simon, H. A. (1996c). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology*, 31(1), 1-40.
- Graesser, A. C., & Clark, L. F. (1985). *Structures and procedures of implicit knowledge*. Norwood, NJ: Ablex.
- Groen, G. J., & Patel, V. L. (1988). The relationship between comprehension and reasoning in medical expertise. In M. T. H. Chi, R. Glaser, & M. Farr (Eds.), *The nature of expertise*, (pp. 287-310). Hillsdale, NJ: Erlbaum.
- Jensen, A.R. (1990). Speed of information processing in a calculating prodigy. *Intelligence*, 14, 259-274.
- Joseph, G. M., & Patel, V. L. (1990). Domain knowledge and hypothesis generation in diagnostic reasoning. *Journal of Medical Decision Making*, 10, 31-46.
- Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, NJ: Erlbaum.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95(2), 163-182.
- Koedinger, K. R., & Anderson, J. R. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science*, 14, 511-550.
- Lemieux, M., & Bordage, G. (1992). Propositional versus structural semantic analyses of medical diagnostic thinking. *Cognitive Science*, 16, 185-204.
- McKeithen, K. B., Reitman, J. S., Rueter, H. H., & Hirtle, S. C. (1981). Knowledge organization and skill differences in computer programmers. *Cognitive Psychology*, 13, 307-325.
- Patel, V. L., Arocha, J. F., & Groen, G. J. (1986). Strategy selection and degree of expertise in medical reasoning. In C. Clifton (Ed.), *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*, (pp. 780-791). Hillsdale, NJ: Erlbaum.
- Patel, V. L., Evans, D. A., & Groen, G. J. (1989). Biomedical knowledge and clinical reasoning. In D. A. Evans & V. L. Patel (Eds.), *Cognitive science in medicine: Biomedical modeling*, (pp. 53-112). Cambridge, MA: MIT Press.
- Patel, V. L., & Groen, G. J. (1986). Knowledge based solution strategies in medical reasoning. *Cognitive Science*, 10, 91-116.
- Patel, V. L., & Groen, G. J. (1991a). The general and specific nature of medical expertise: A critical look. In A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits*, (pp. 93-125). New York: Cambridge University Press.
- Patel, V. L., & Groen, G. J. (1991b). Developmental accounts of the transition from medical student to doctor: Some problems and suggestions. *Medical Education*, 25(6), 527-535.
- Patel, V. L., Groen, G. J., & Arocha, J. F. (1990). Medical expertise as a function of task difficulty. *Memory and Cognition*, 18(4), 394-406.
- Patel, V. L., & Kaufman, D. R. (1995). Clinical reasoning and biomedical knowledge: Implications for teaching. In J. Higgs & M. Jones (Eds.), *Clinical reasoning in the health professions*, (pp. 117-128). Oxford, UK: Butterworth Heinemann.
- Patel, V.L., & Medley-Mark, V. (1986). Relationship between representation of textual information and underlying problem representation in medicine. *Cognitive research series*. Montreal: McGill University, Centre for Medical Education. (CME86-CS1).
- Reitman, J. S. (1976). Skilled perception in Go: Deducing memory structures from inter-response times. *Cognitive Psychology*, 8, 336-356.
- Rikers, R. M. J. P., Schmidt, H. G., & Boshuizen, H. P. A. (1997, March). *Nonroutine problem solving by medical experts*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, ILL.

- Schmidt, H. G., & Boshuizen, H. P. A. (1992). Encapsulation of biomedical knowledge. In D. A. Evans & V. L. Patel (Eds.), *Advanced models of cognition for medical training and practice*. New York, NY: Springer Verlag.
- Schmidt, H. G., & Boshuizen, H. P. A. (1993). On the origin of intermediate effects in clinical case recall. *Memory and Cognition*, 21(3), 338-351.
- Schmidt, H. G., Boshuizen, H. P. A., & Hobus, P. P. M. (1988). Transitory stages in the development of medical expertise: The "intermediate effect" in clinical case representation studies, *Proceedings of the Cognitive Science Society Meeting*, (pp. 139-145). Hillsdale, NJ: Lawrence Erlbaum.
- Schmidt, H. G., Boshuizen, H. P. A., & Norman, G. R. (1992). Reflections on the nature of expertise in medicine. In E. Keravnou (Ed.), *Deep models for medical knowledge engineering*. Amsterdam, The Netherlands: Elsevier.
- Shanteau, J. (1992). How much information does an expert use? Is it relevant? *Acta Psychologica*, 81(1), 75-86.
- Simpson, S. A., & Gilhooly, K. J. (1997). Diagnostic thinking processes: Evidence from a constructive interaction study of electrocardiogram (ECG) interpretation. *Applied Cognitive Psychology*, 11(6), 543-554.
- Thomas, C. L. (1989). *Taber's Cyclopedic Medical Dictionary*. (16th ed.) Philadelphia : Davis, cop.
- Van de Wiel, M.W.J., Boshuizen, H.P.A., & Schmidt, H.G. (1998). A failure to reproduce the intermediate effect in clinical case recall. *Academic Medicine*, 73, 894-900.
- Van Dijk, T. A. & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Vicente, K. J., & Wang, J. H. (1998). An ecological theory of expertise effects in memory recall, *Psychological Review*, 105(1), 33-57.
- Voss, J.F., & Bisanz, G. L. (1985). Knowledge and the processing of narrative and expository texts. In B. K. Britton & J. B. Black (Eds.), *Understanding of expository text* (pp. 173-198). Hillsdale, NJ: Erlbaum.

Appendix 4.1

Case of heart failure

A 45-year old woman has complaints of increasing dyspnea and ankle edema since six weeks. The patient has always felt a fast heartbeat.

History: Epileptic from the age of fifteen. Family history: Negative. Social history: She owns a homeopathic shop. Intoxications: She smokes 20 cigarettes a day. Medication: Depakine 3x300 mg.

Physical examination:

Physical examination revealed a thin, dyspnoeic woman. Her weight was 50 kg., and her height 1.70 m. Pulse rate was 100/min and regular. Blood pressure was 140/90 mmHg. Carotid artery pulses normal. Distended jugular veins. Thyroid gland was enlarged and elastic on palpation. Heart: Apical impulse was just lateral to the midclavicular line. First and second heart sounds are normal. Left and right ventricular pulsations. There was a 3/6 holosystolic murmur at the apex. Lungs: Normal breath sounds, basal rales and dullness on percussion. Abdomen: The liver was enlarged (four fingers). Extremities: Pronounced bilateral ankle edema.

Laboratory findings:

ESR 6 mm/h (normal: <12 mm/h), Hb 9.1 mmol/l (normal: 7.4-9.6 mmol/l), Ht 0.42 l (normal: 0.36-0.46 l), WBC 6.8 10⁹/l (normal: 3.4-8.9 10⁹/l), Platelet count 153 10⁹/l (normal: 150-350 10⁹/l), PTT 31 sec (normal: 24-35 sec), Sodium 140 mmol/l (normal: 132-142 mmol/l), Potassium 5.25 mmol/l (normal: 3.6-5.0 mmol/l), Calcium 2.27 mmol/l (normal: 2.1-2.6 mmol/l), Urea 4.5 mmol/l (normal: 3.0-7.0 mmol/l), Creatinine 67 mmol/l (normal: 53-97 mmol/l), Alkaline phosphatase 129 U/l (normal: 30-125 U/l), GGT 42 U/l (normal: < 50 U/l), ASAT 27 U/l (normal: < 35 U/l), ALAT 19 U/l (normal: < 35 U/l), CPK 86 U/l (normal: 40-200 U/l), LDH 381 U/l (normal: 200-450 U/l), Total protein 68 g/l (normal: 65-79 g/l), Albumin 36 g/l (normal: 40-60 g/l),
Arterial blood gases: pH 7.41 (normal: 7.35-7.45), pCO₂ 5.1 kPa (normal: 4.5-5.9 kPa), pO₂ 10.5 kPa (normal: 8.7-13.1 kPa), O₂ saturation 94% (normal: 93-98 %).

Electrocardiogram:

Sinus rhythm, rate 90/min. Left atrial enlargement. Intermediate position of the electrical axis. Poor R wave progression in the left precordial leads. Abnormal ST-T segments.

Chest X-ray:

Cardiothoracic ratio > 0.5. Bilateral pleural effusion (right > left).

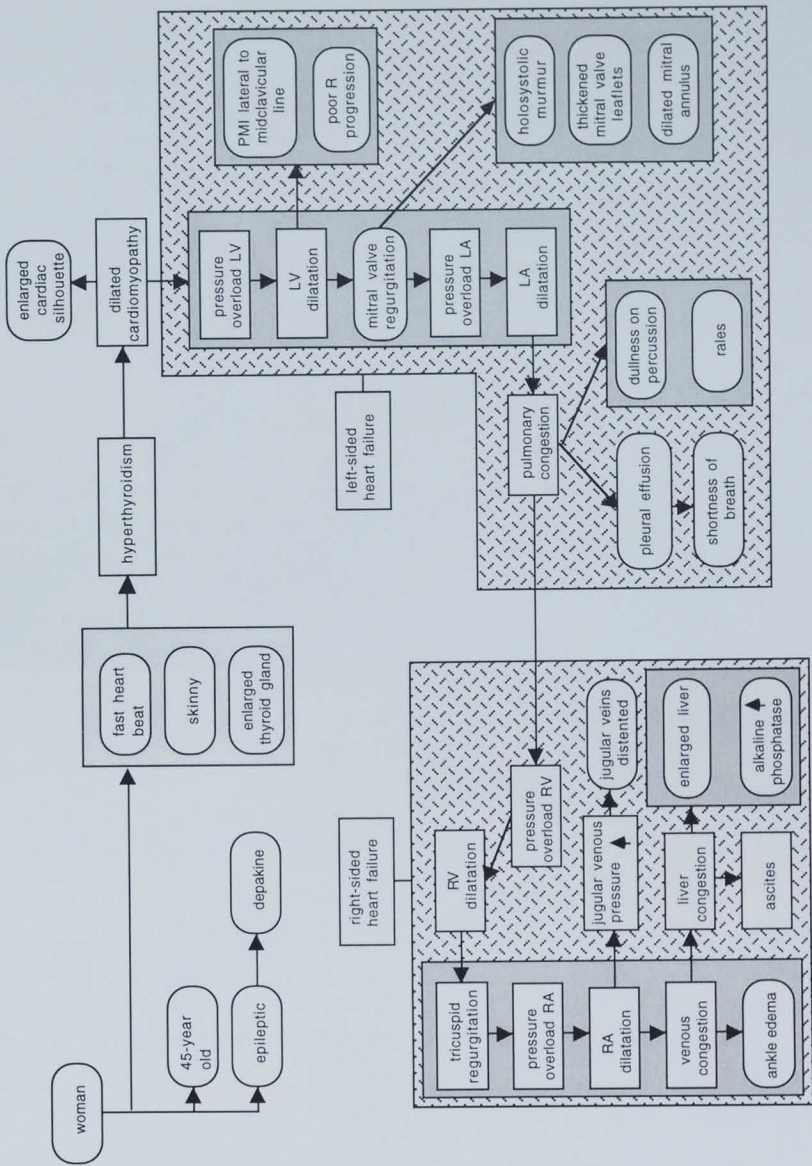
Echocardiography/Doppler:

Global hypokinesia of the dilated left ventricle and severe mitral insufficiency. Mitral valvular annulus enlarged 4 cm. The mitral valve leaflets are slightly thickened.

Appendix 4.2

Canonical Explanation for the Congestive Heart Failure Case

Frames with rounded corners represent signs and symptoms present in the case, and frames with rectangular corners represent concepts that were not present. Most links between concepts in this model indicate a causal relationship.



Chapter 5

Effects of Clinical Case Priming on the Activation of Encapsulated Knowledge: Differences between Medical Experts and Subexperts¹

This paper addresses the role of so-called encapsulated knowledge in clinical case representations by advanced medical students, medical subexperts (i.e., expert physicians diagnosing clinical cases outside their specialty), and expert physicians. That is, to what extent is encapsulated knowledge used within and outside the expert physicians' domain of expertise? This question is addressed in the first experiment with fourth-year medical students, subexperts from six different specialties, and internists (the expert group), who studied two clinical cases within the domain of internal medicine. After each case, participants had to evaluate three types of medical items (inferred encapsulating items, literally stated signs and symptoms, and filler items) whether or not they were related to the clinical case. The results of the experiment showed that experts, in contrast to both other groups, were faster and more accurate in evaluating the relatedness of the types of items. In particular, the experts' performance on encapsulating items led to fast and accurate responses. The second experiment involved fourth-year medical students, subexperts, and cardiologists (the experts) who had to study a text containing the most important clinical manifestations associated with a disease called congestive heart failure. After they had studied the text, participants were instructed to evaluate three types of items: Biomedical items, encapsulating items, and filler items. Participants had to indicate as fast as possible whether the items were related to the text. Results showed that experts outperform students and subexperts on the different item types. Again, encapsulating items led to the fastest and most accurate responses. Findings of both experiments indicate that especially encapsulated knowledge plays an important role in the expert physicians' clinical case representation within and outside their specialty.

¹ Parts of this chapter appeared as Rikers, R.M.J.P., Schmidt, H.G., Boshuizen, H.P.A. (1999, April). *Clinical case representations by medical experts and subexperts*. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal.

Introduction

A distinctive and stable criterion of medical expertise is the high accuracy of expert physicians diagnosing clinical cases within their domain of expertise (e.g., Boshuizen, 1989; Claessen & Boshuizen, 1985; Patel & Groen, 1986; Rikers, Boshuizen, & Schmidt, 1997; Schmidt & Boshuizen, 1993). Early attempts to explain the superior performance of expert physicians primarily focused on the strategies experts used while solving diagnostic problems (e.g., Elstein, Schulman, & Sprafka, 1978). This line of research, however, only found small differences between expert physicians, (advanced) medical students, and novices. More recent, and more successful approaches, shifted their attention towards the structure of medical knowledge (e.g., Bordage, 1994; Custers, Boshuizen, & Schmidt, 1996; Feltovich & Barrows, 1984; Schmidt & Boshuizen, 1992, 1993). What these studies have shown is that an important transition takes place in the way knowledge is organized. That is, medical students do not only know less about medicine than expert physicians do, but more importantly, they also apply qualitatively different knowledge while diagnosing a clinical case.

According to Schmidt and Boshuizen (1992, 1993) students' knowledge structure is best characterized by the prevalence of biomedical knowledge (i.e., knowledge concerning the mechanisms or processes underlying the manifestations of a disease), which they acquired during the first years of medical school. As competence is attained, elements of their biomedical knowledge become increasingly interconnected, so that it can be accessed as coherent 'chunks' of information rather than fragments. At this stage, shortcuts in their biomedical reasoning begin to emerge (cf. Blessing & Anderson, 1996; Koedinger & Anderson, 1990), and it is not necessary anymore to activate all possible relevant knowledge in order to understand the signs and symptoms; mainly knowledge that is pertinent for understanding will be activated. By the time they have reached the level of an expert physician, their original elaborate biomedical knowledge structure has changed into a highly condensed causal representation. That is, as a result of extensive practice and the confrontation with actual patients their biomedical knowledge has become linked with or *encapsulated* under a limited number of clinically relevant concepts. Therefore, expert physicians that diagnose a clinical case will mainly activate specific pre-existing encapsulating concepts. For instance, if one would require an expert physician to study a clinical case description that contains, among other things, the following information: *A 45-year old woman who has always felt a fast heartbeat. Her weight is 50 kg., and her height 1.70 m. Her thyroid gland is enlarged and elastic on palpation.* The expert's representation based on this information will most likely contain the concept of *hyperthyroidism*, which is an (knowledge-based) inference from the text that summarizes or encapsulates the signs and symptoms displayed

(cf. Graesser, Singer, & Trabasso, 1994). Medical students, on the other hand, will have considerable difficulty to establish a meaningful relationship between the signs and symptoms, and their representation will most likely not contain a concept such as hyperthyroidism. Instead, students will almost literally reproduce the information as it is given without inferences as made by the expert physician.

Evidence to support encapsulation theory has been repeatedly found in clinical case studies. In these studies participants of different levels of expertise were instructed to study a text describing a patient's history, complaint(s) and physical condition. In addition, they had to provide a diagnosis, write down everything they remembered of the case (i.e., free recall), and finally they had to explain the signs and symptoms in the case (i.e., pathophysiological explanation). The outcome of these studies was that advanced students (i.e., intermediates) were able to recall more information of the text and explained the described sign and symptoms in a more elaborate fashion than did expert physicians. So, rather than a monotonically increasing function with level of expertise, an inverted U-shaped relation was found (e.g., Boshuizen, 1989; Coughlin & Patel, 1986; Patel & Medley-Mark, 1986; Schmidt & Boshuizen, 1993). This phenomenon that has become known as the *intermediate effect* can easily be explained by encapsulation theory. In contrast to (advanced) students, expert physicians' recall and pathophysiological protocols contain a small number of highly inclusive concepts. Students, on the other hand, will almost literally reproduce the stated information and provide deep and elaborate explanations for the signs and symptoms of the case. Consequently, they remember and explain "more" than do expert physicians. Encapsulation of the physicians' biomedical knowledge does not imply that their biomedical knowledge is lost or has become less accessible. On the contrary, physicians can easily retrieve their biomedical knowledge if they need it. For example, if they are confronted with complicated cases, or cases that they have not frequently encountered (Norman, Trott, Brooks, & Smith, 1994; Patel, Groen, & Arocha, 1990). That is, their encapsulated knowledge can, at any moment, be "decapsulated" when it fails in establishing a meaningful representation of a clinical case (Schmidt & Boshuizen, 1993; Schmidt, Boshuizen, & Norman, 1992).

However, previous clinical case studies comparing medical experts with subexperts (i.e., medical experts confronted with problems outside their specialty), showed ambiguous results concerning the role of encapsulated knowledge within and outside the physician's domain of expertise (e.g., Patel, Arocha, & Groen, 1986, 1990; Rikers, Boshuizen, & Schmidt, 1997; Rikers, Schmidt, & Boshuizen, in press). For instance, in the study by Rikers, Boshuizen and Schmidt (1997), experts and subexperts had to diagnose, recall and explain a clinical case description. It was predicted that if encapsulated knowledge is confined to the physicians' domain of expertise, subexperts

would not be able, or have considerably difficulty to apply encapsulated knowledge. Surprisingly, the measures of free recall and pathophysiology failed to unearth significant differences (i.e., experts and subexperts did not significantly differ in the total number of propositions and encapsulating concepts in recall and pathophysiology). Findings that were explained by assuming that subexperts had enough *generic expertise* to retain and explain the relevant information presented in the clinical case (Patel & Groen, 1991; Patel, Arocha, & Groen, 1990). Consequently, the traditional measures of free recall and pathophysiology might be appropriate if experts and students are compared, but they are not eligible candidates to reveal differences if experts and subexperts are under investigation.

In order to overcome these difficulties associated with investigating clinical case representations of medical experts and subexperts, two experiments are reported introducing the so-called "priming paradigm" in clinical case studies. That is, instead of recalling or explaining the information of a clinical case description, participants have to decide as fast as possible whether a medical concept presented on a computer screen is related to the clinical case. The clinical case functions as a kind of "prime" for the concepts participants activate during case processing. The resulting pre-activation of concepts is assumed to facilitate the subsequent evaluation process that requires access to this information in memory, thereby producing a *semantic-priming effect* (Collins & Loftus, 1975).

Experiment 1

In Experiment 1 the role of encapsulated knowledge is explored within and outside the expert physicians' domain of expertise, avoiding the "traditional" approach of free recall and pathophysiological explanations. Internists (the expert group), subexperts (from six different specialties), and fourth-year medical students, were instructed to study clinical cases that were within the domain of internal medicine. For the internists, the clinical cases were considered to be well-known domain problems. Subsequently, the participants had to decide, as fast as possible, whether the items presented on a computer screen were related to the clinical case they had studied. Three types of items were used: Literally stated *signs and symptoms*, inferred *encapsulating items* that were not part of the text, and *filler items* that were unrelated to the case. Based on the model of knowledge encapsulation it is expected that sign and symptoms displayed in a clinical case activate specific pre-existing encapsulating concepts of the expert physician. Medical students and subexperts, on the other hand, will not be able, or have considerable difficulty, to link the signs and symptoms to pre-existing encapsulating concepts, and consequently, their encapsulated knowledge will be activated to a lesser extent, or

not at all. In order to avoid the possibility that experts, but especially subexperts and medical students fail to activate important case-related pre-existing concepts (because they completely or partly missed the correct diagnosis), one half of all participants received the correct diagnosis of each case in advance. Providing the correct diagnosis enables students, subexperts, and experts equally to activate whatever prior knowledge they think is essential to understand the clinical case.

Method

Participants

Participants were 22 fourth-year medical students and 44 medical specialists of the Maastricht University Hospital: 22 internists (the experts) and 22 physicians from six different specialties (the subexperts). All physicians were practitioners with at least MD degrees and Board Certification in their respective specialties and had at least 6 years of experience. The participants received a financial compensation for their participation.

Material

The materials consisted of two clinical case descriptions presented on a computer screen. Depending on the experimental condition, half of the participants got the correct diagnosis of the cases in advance. Both cases were within the domain of internal medicine and reported some contextual information of the patient, the complaint(s), findings from history taking and physical examination, relevant laboratory data and some additional findings. One case describes a patient with acute bacterial endocarditis (Patel & Groen, 1986), the other a patient with heart failure (Van de Wiel, Boshuizen, Schmidt, & de Leeuw, 1993). The cases were considered to be non-routine problems for the subexperts and routine problems for the internists. The case descriptions were about one page in length and consisted of 49 and 43 propositions respectively. In addition, 32 items per case were assembled: 8 Encapsulating items, 8 Signs and Symptoms, and finally 16 fillers that were not related to the clinical case (see Appendix 5.1). As has been outlined in the introduction, encapsulating items are high-level inferences (i.e., based on at least two propositions in the text). In order to identify a priori, potential, high-level inferences, verbal (i.e., think-aloud) and written protocol data from previous studies using exactly the same cases were assembled (Schmidt & Boshuizen, 1992, 1993; Van de Wiel, Boshuizen, Schmidt, & de Leeuw, 1993). The most common encapsulating items were selected. The signs and symptoms that

were used had been explicitly stated in the text. The correct response to both encapsulating items and signs and symptoms was *yes*.

Procedure

The participants were instructed to carefully study each clinical case for a period of one minute. Subsequently, they had to evaluate items on the basis of their relatedness to the clinical case. That is, they had to respond to the question: "Is this item related to the clinical case"? All items were randomly presented on the computer screen. A row of asterisks appeared in the center of the screen for 500 ms and was replaced after 500 ms by an item, centered on the screen, which remained visible until the participant made a response. Participants were instructed to respond as quickly as possible to each item without sacrificing accuracy by pressing the "/" key for *yes* and the "z" key for *no*. Once a response was made, the item was removed and a blank screen remained for 500 ms, after which the row asterisks reappeared marking the start of the next trial. The computer automatically registered response times after the key was pressed. Each participant had to evaluate both cases. Participants who did not get the correct diagnoses in advance (i.e., the "unstated condition"), had to provide a diagnosis for both cases after they had completed all other tasks. Carry-over effects were controlled by balancing the presentation order of the cases. To enable the participants to get used to their task, the experimental cases were preceded by one sample case.

Analysis

The data were analyzed using ANOVAs, with Expertise Level (Students, Subexperts, and Experts) and Diagnosis (Stated and Unstated) as the between-subjects factors and Item Type (Encapsulated and Signs & Symptoms) as the within-subjects factor (significance level was set at .05). Fillers were excluded from the analysis because there were no significant differences found between expertise levels. For each participant, the data from the two clinical cases were averaged to obtain a mean response time and error rate at both levels of Item Type. Subsequent analysis of the response times included data from all trials, regardless of whether responses were correct or incorrect. This approach was taken because there would otherwise be too few correct responses at student level to analyze. Including only response times from correct trials, however, does not significantly alter the pattern of results. The accuracy of the provided diagnosis of the participants in the Unstated condition was rated by two independent domain experts on a 7-point scale from 0 (completely incorrect diagnosis) to 6 (completely correct diagnosis). If an accurate diagnosis was given as part of a differential list of diagnoses, it was

considered accurate. Inter-rater agreement was higher than 95 %, and differences were resolved through discussion.

Results

Diagnostic Accuracy

Figure 5.1 shows the mean diagnostic accuracy of all diagnoses produced by experts, subexperts, and medical students in the diagnostically unstated condition. Analysis of variance showed a significant main effect for Expertise Level [$F(2,30) = 21.22$, $MSE = .292$, $p < .0001$]. Newman-Keuls test for diagnoses showed that internists were more accurate than both other groups. The difference between students and subexperts was not significant. It is important to note that none of the participants provided a completely accurate diagnosis on either one of the cases.

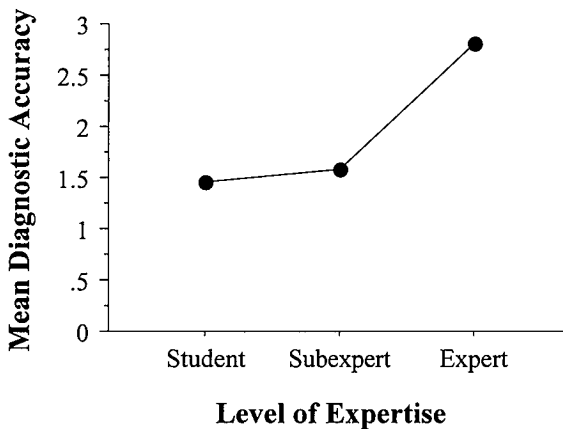


Figure 5.1. Mean diagnostic accuracy as a function of expertise level.

Response Times

The mean response times of experts, subexperts, and students for encapsulating items and signs & symptoms are shown in Figure 5.2. Analysis of variance of the mean response times shows a main effect of Expertise Level [$F(2,60) = 7.81$, $MSE = .294$, $p < .001$], Diagnosis [$F(1,60) = 19.03$, $MSE = .294$, $p < .0001$], but no main effect of Item Type [$F(1,60) = .041$, $MSE = .049$]. A significant interaction between Item Type x Diagnosis is found [$F(2,60) = 5.41$, $MSE = .049$, $p < .01$]. There were no other significant two or three-way interac-

tions: Expertise Level x Diagnosis [$F(2,60) = .539$, $MSE = .294$]; Item Type x Diagnosis [$F(1,60) = .189$, $MSE = .049$]; Item Type x Expertise Level x Diagnosis [$F(2,60) = .262$, $MSE = .049$].

Expert physicians were faster in evaluating both types of items than any other group. Especially encapsulated items led to speeded responses of the expert physicians. However, students and subexperts benefited a little bit more than expert physicians did if they were informed about the correct diagnoses. Roughly, students and subexperts were 25% faster and experts were 20% faster. Although the basic pattern of response times for both types of items does not seem to alter, participants were considerably faster if a diagnosis was provided. Interestingly, this is also true for the expert physicians.

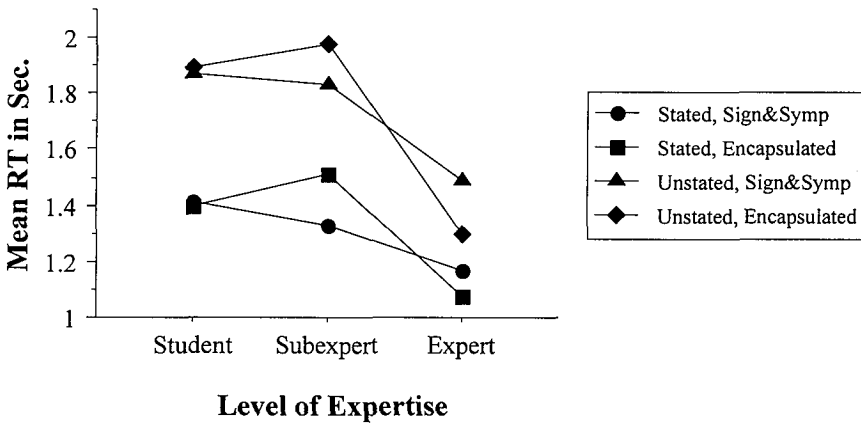


Figure 5.2. Mean response times as a function of expertise level, diagnosis, and item type.

Error Rates

Figure 5.3 shows the mean error rates of experts, subexperts, and students for both types of items. Analysis of variance of the mean error rates shows a main effect of Expertise Level [$F(2,60) = 20.99$, $MSE = 1.261$, $p < .0001$], and Item Type [$F(1,60) = 23.67$, $MSE = .933$, $p < .0001$]. No main effect of Diagnosis was found [$F(1,60) = .865$, $MSE = 1.261$]. A significant interaction between Item Type x Diagnosis is found [$F(2,60) = 13.36$, $MSE = .933$, $p < .001$]. There were no other significant two or three-way interactions: Expertise Level x Diagnosis [$F(2,60) = 1.947$, $MSE = 1.261$]; Item Type x Expertise Level [$F(2,60) = .887$, $MSE = .993$]; Item Type x Expertise Level x Diagnosis [$F(2,60) = 1.56$, $MSE = .933$].

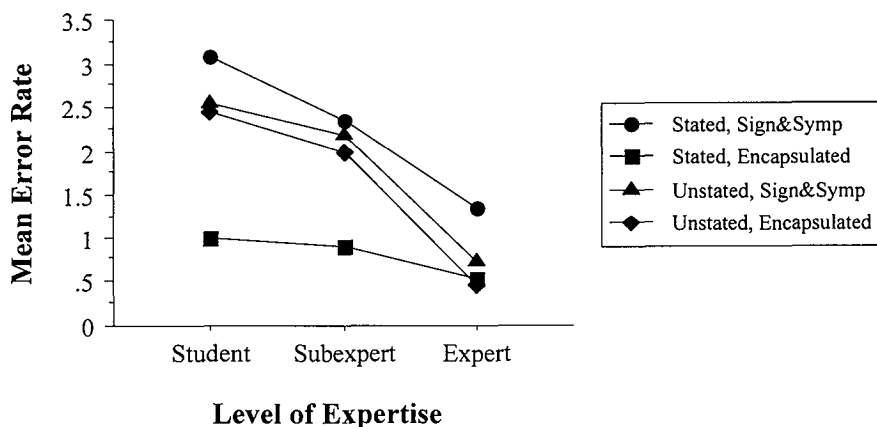


Figure 5.3. Mean error rates as a function of expertise level, diagnosis, and item type.

Discussion

The results of experiment 1 are consistent with the hypotheses derived from encapsulation theory as proposed by Schmidt and Boshuizen (1992, 1993). It was predicted that expert physicians, diagnosing clinical cases within their domain of expertise, would mainly rely on their pre-existing encapsulated knowledge in order to build a suitable representation of a clinical case. Subexperts and medical students, on the other hand, will have difficulty to link signs and symptoms to encapsulated knowledge. Findings corroborate these predictions: Experts did not only achieve a higher diagnostic accuracy, they were also faster and more accurate in evaluating case related items than did subexperts and medical students. Findings also indicate that encapsulated knowledge plays a more important role within the expert physician's case representation than literally stated signs and symptoms. The opposite seems true for the subexperts: They need more time to evaluate encapsulating items than literally stated signs and symptoms. However, the clear distinction between experts and subexperts in their processing speed of encapsulating items and signs and symptoms does not seem to exist for students. Neither encapsulating items nor signs and symptoms seem to have a special status in their knowledge structure.

Furthermore, students, subexperts, and experts are approximately 400 ms faster on both types of items if they are informed about the correct diagnoses, then when they are not. The availability of the accurate diagnosis seems to "prime" the encapsulating items and signs and symptoms. Interestingly, experts also seem to benefit almost equally from the availability of the correct

diagnoses. This finding can be explained by the diagnostic performance of the expert physicians. A mean diagnostic accuracy of 50% of the maximum score indicates that, although experts outperformed both other groups, they did not detect all essential components of the correct diagnoses.

The differences in response times between experts, subexperts, and students may be explained by a different approach to evaluate the items. The experts' duration of response times might be an indication that they had to retrieve the encapsulating items from long-term memory or long-term working memory (LTWM) and were therefore no part of their active representation in working memory while performing the decision task (Ericsson & Kintsch, 1995). Response times of subexperts and students, on the other hand, might be an indication that they engaged in (slow) strategic processes, constructing inferences not encoded in the mental representation of the case during its original reading. In other words, experts activated pre-existing inferences during the course of comprehension (i.e., on-line), subexperts, and students during the later decision task (i.e., off-line). However, it is important to note that based on the present methodology we cannot exclude the possibility that expert physicians also generated their inferences off-line (Graesser & Bower, 1990; Graesser, Singer, & Trabasso, 1994; McKoon & Ratcliff, 1992). Further research is necessary to disentangle this issue.

Interestingly, expert physicians were not only faster in evaluating encapsulating items, they were also faster in evaluating literally stated sign and symptoms. Items such as "Shaking Chills," "Sweating," "Hypotension," "Rales," and "Temperature 40° C," are surely not restricted to one medical specialty. Therefore, it is surprising that our subexperts, all board certified medical specialists with many years of experience, were considerably slower on these items than the experts. The model of knowledge encapsulation might explain this finding. Signs and symptoms reported in a clinical case point towards specific pre-existing encapsulating items in LTM. This results in highly activated encapsulating concepts within the expert physician's case representation. Signs and symptoms on their turn benefit from these high activation levels of the encapsulating concepts, because some of this activation flows back to them.

Experiment 2

Findings of Experiment 1 corroborated the central position of encapsulated knowledge within the expert physician's clinical case representation. Experts were faster and more accurate in evaluating encapsulating items than did medical students and subexperts. However, the status of biomedical knowledge within the expert physician's clinical case representation remained unclear. Therefore, the aim of Experiment 2 is to investigate the role of biomedical knowledge in the clinical case representation of expert physicians. Previous studies, mainly using a post-hoc methodology, do not provide a sufficient basis to decide whether biomedical knowledge plays a major role in the expert physicians' representation (e.g., Lesgold, 1984; Lesgold, Robinson, Feltovich, Glaser, Klopfer, & Wang, 1988; Patel, Evans, & Chawla, 1986; Patel & Groen, 1986; Patel, Groen, & Arocha, 1990). Based on these studies, three positions on the role of biomedical knowledge can be distinguished: (1) biomedical knowledge becomes *prevalent* in the course towards expertise; (2) biomedical knowledge becomes *inert*; (3) biomedical knowledge becomes *encapsulated*.² Evidence to support the position that biomedical knowledge becomes prevalent is provided by the studies by Lesgold and colleagues (1984, 1988). Their studies provide evidence for the view that the development of medical expertise is characterized by an increased and flexible use of biomedical knowledge. Patel and colleagues (Patel & Groen, 1991; Patel, Evans, & Groen, 1989; Patel & Kaufman, 1995), on the other hand, take an opposite position. They assume that biomedical knowledge becomes "inert" in the course towards expertise, and is therefore not used by expert physicians in their clinical case representation. Instead, clinicians use their clinical knowledge in order to build an adequate case representation. Finally, a third position on the debate of the role of biomedical knowledge is expressed by Schmidt and Boshuizen (1992, 1993). They assume that biomedical knowledge is still available, but it plays no longer a prominent role in the expert physician's case representation. That is, during the development of medical expertise, biomedical knowledge has become encapsulated into clinically relevant concepts. For instance, in one of their studies, Schmidt and Boshuizen (1993) instructed participants of different levels of expertise to activate their pathophysiological knowledge of endocarditis for either 3.5 min. or 30 sec. After completing this activation task, participants were requested to study a clinical case description of acute bacterial endocarditis for 30 sec. Subsequently, they were asked to write down whatever information they recalled from the case and to state the most likely diagnosis. It was predicted that, among novices

² It is important to note that a fourth position might be taken in this discussion on the role of biomedical knowledge. This position implies that biomedical knowledge is acquired and applied in the early stages of the development of expertise, but will become gradually less important or even inaccessible. However, empirical evidence to support this position is not very strong (see for a more detailed discussion of the different positions, Boshuizen & Schmidt, 1992).

and advanced students the amount of relevant knowledge activated by the priming procedure would influence the comprehension of the clinical case and, hence, its recall. The outcome of this study was - as predicted - that the priming of prior pathophysiological knowledge had its largest impact on case processing by the students of intermediate levels of expertise, whereas the experts' performance were not affected. That is, experts produced *more* pathophysiological information at both priming intervals, but the experts' recall was unaffected by the degree of priming. These data provide unequivocal evidence for the assumption that intermediates' comprehension of a clinical case is based on detailed pathophysiological knowledge, while experts' understanding is independent of elaborate processing of this kind. Furthermore, this study also showed that expert physicians are still able to apply their biomedical knowledge if task requirements ask them to do so. However, although this study strongly suggest that experts' clinical case representation is built upon encapsulated knowledge, it is not possible to assess the role of biomedical knowledge within the representation. That is, do expert physicians activate biomedical knowledge at all during case processing, or is their representation solely build on encapsulated knowledge?

Based on the different perspectives on the role of biomedical knowledge, the present study investigates the role of biomedical knowledge in the clinical case representations by medical experts, subexperts, and fourth-year medical students. In line with Experiment 1, participants had to evaluate three item types: Biomedical items, encapsulating items, and filler items that were not related to the case. It is predicted that if biomedical knowledge becomes encapsulated, expert physicians will show speeded responses evaluating encapsulating items. Biomedical items, on the other hand, will take considerably more time to evaluate, because they do not take such a central position within the expert physician's case representation. Subexperts and students, who have to resort to biomedical knowledge will show the opposite preference: Their clinical case representation will, first of all, consists of biomedical items, and therefore encapsulating items will take more time to evaluate.

Method

Participants

Participants were 9 fourth-year medical students and 18 medical specialists of the Maastricht University Hospital. The specialists were 9 cardiologists (the domain experts in this study) and 9 physicians from six different specialties (the subexperts). All physicians were practitioners and had at least 10 years of experience. The medical specialists had at least MD degrees and Board

Certification in their respective specialties. The participants received a financial compensation for their participation.

Material

The materials were presented on a computer screen (Macintosh PowerBook) and consisted of a text describing the clinical manifestations associated with congestive heart failure. The description was about one page in length and was a 575-word exposition entitled "Congestive Heart Failure." The facts included all typical sign and symptoms related to congestive heart failure (e.g., breathlessness, abdominal discomfort, edema and weakness). Statements related to heart failure's pathophysiology (i.e., statements related to the disturbed underlying processes) were avoided as much as possible (see Appendix 5.3). Based on the pathophysiological explanations of previous clinical case studies with case of congestive heart failure, 60 items were assembled: 20 Encapsulating items, 20 biomedical items, and finally 20 filler items that were not related to the description (see Appendix 5.2). Both encapsulating and biomedical items were not stated in the text. The correct response to the encapsulating items and biomedical items was *yes*.

Procedure

Participants were instructed to read a text describing the clinical manifestations associated with congestive heart failure for a period of three minutes. They were instructed to study the text with the aim to account for all sign and symptoms. Finally, they had to evaluate items on the basis of their relatedness to the clinical case. That is, they had to respond to the question: "Is this item related to the text"? All items were randomly presented on a computer screen. A row of asterisks appeared in the center of the screen for 500 ms and was replaced after 500 ms by an item, centered on the screen that remained visible until the participant made a response. Participants were instructed to respond as quickly as possible to each item without sacrificing accuracy by pressing the "/" key for *yes* and the "z" key for *no*. Once a response was made, the item was removed and a blank screen remained for 500 ms, after which a row asterisks appeared marking the start of the next trial. The clinical text was preceded by one sample text in order to enable the participants to get used to their task.

Analysis

The data were analyzed using repeated measures ANOVAs with Expertise Level as the between-subjects factor and Item Type as the within-subjects factor. Analysis of the response times included data from all trials, regardless

of whether responses were correct or incorrect. Including only response times from correct trials does not alter the pattern of results. In addition, the Student-Newman-Keuls test was used to make pairwise comparisons between the different levels of expertise (significance level was set at .05).

Results & Discussion

Response Times

The computer registered the time each participant needed to make a response. The mean response times of the experts and subexperts for encapsulating items and biomedical items are shown in Figure 5.4. Analysis of variance of the mean response times indicated a significant difference for Expertise Level [$F(2,24) = 3.68, MSE = 1.084, p < .05$] and a significant main effect of Item Type [$F(1,24) = 36.27, MSE = .064, p < .0001$]. Interaction between Expertise level and Item Type was not significant [$F(2,24) = 1.62, MSE = .064, p > .20$]. Newman-Keuls tests for Item Type showed significant differences between internists and both other groups. Students and Subexperts, on the other hand, did not show significant differences.

Expert physicians evaluated both item types faster than did subexperts and students, as was predicted based on encapsulation theory. Similar to the findings of Experiment 1, especially encapsulating items led to speeded responses of the expert physicians. However, this was not only the case for the expert physicians, subexperts and students also showed a speed-up in response times for the encapsulating items. Furthermore, the absence of significant differences in response times between

Subexperts and students suggest that subexperts did relapse into an elaborate processing approach while processing the clinical case descriptions.

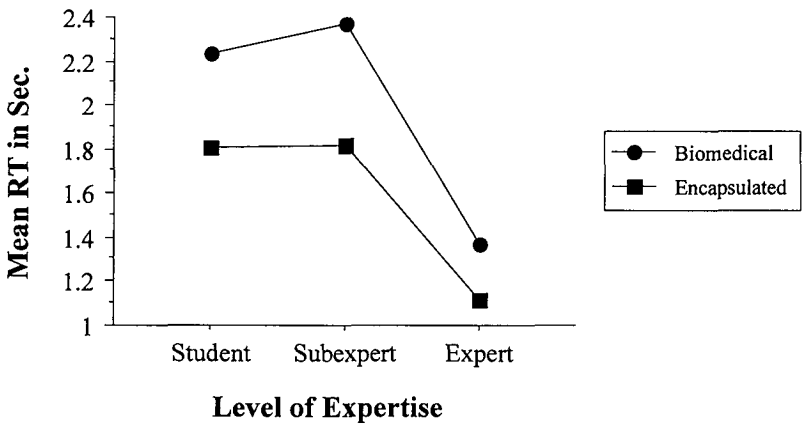


Figure 5.4. Mean response times as a function of expertise level and item type.

Error Rates

Figure 5.5 shows the mean error rates of experts, subexperts, and medical students for both item types. Analysis of variance of the error rates showed no main effects of Expertise Level and Item Type. Finally, no interaction between Expertise Level and Item Type was found [all F 's < 1]. In contrast to Experiment 1, no significant differences between item types and expertise levels were found. Note that error rates of all expertise levels are not very high.

This finding, and the fact that students and subexperts also needed more time to evaluate biomedical items, seems to contradict the central notion of encapsulation theory that biomedical items are central within the students' and subexperts' case representation. A possible explanation for this outcome might be that knowledge encapsulation takes place in an earlier phase than is suggested by encapsulation theory. Furthermore, congestive heart failure has a high morbidity in the Netherlands (and worldwide), and therefore students and subexperts are not unfamiliar with this disease. That is, they might have seen patients with this disease to an extent that encapsulation of their knowledge base related to this disease occurred in an early stage of their expertise development. Nonetheless, encapsulating theory is the only theory of those mentioned above, which gives the most adequate account for the present data. Therefore, findings of Experiment 1 seem to suggest that the main difference between experts and subexperts is not the presence or absence of encapsulated knowledge, as is stated by encapsulation theory, but the speed of which these items become available.

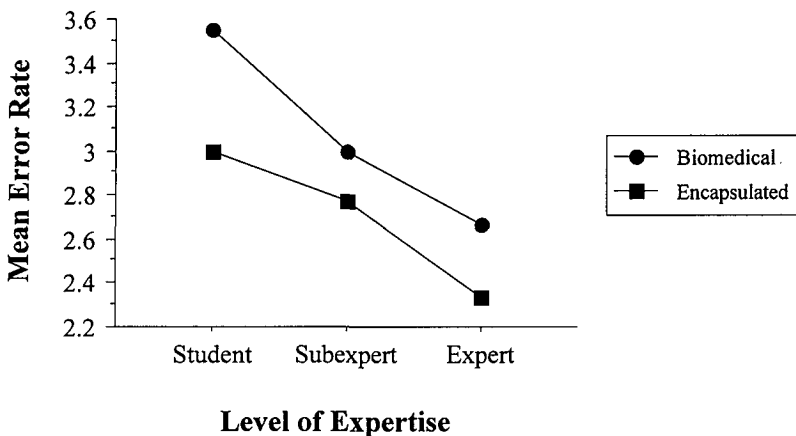


Figure 5.5. Mean error rates as a function of expertise level and item type.

General Discussion

Findings of the present studies demonstrated once more the importance of encapsulated knowledge within the expert physicians' clinical case representation. Experiment 1 demonstrated that experts were not only faster and more accurate in evaluating encapsulating items, literally stated signs and symptoms were also evaluated better by experts than did subexperts and advanced medical students. Moreover, Experiment 2 has shown that encapsulated knowledge also plays a more important role in the expert physicians case representation than biomedical knowledge. Interestingly, this is not only true for expert physicians but also for the subexperts and medical students.

As has been outlined before, within the theory of knowledge encapsulation, encapsulating concepts function as the most important building blocks of the expert physician's clinical case representation. In order to create a coherent representation of a clinical case the expert physician will most likely use relevant pre-existing encapsulating concepts. Biomedical knowledge, on the other hand, comes only into the picture if relevant encapsulating concepts are not available, or if task requirements prescribe them to do so (e.g., Schmidt & Boshuizen, 1993; Van de Wiel, 1997). As has been shown in the present study, this does not necessarily mean that subexperts and students do not make use of encapsulating concepts in their representations. What it does mean is that encapsulating concepts are better integrated within the clinical case representation of an expert physician than that of a subexpert or student.

It is important to note that this study sheds a new light on studies who found no differences between experts and subexperts in recall and pathophysiological explanations (Patel, Arocha, & Groen, 1990; Rikers, Boshuizen, & Schmidt, 1997). For instance, in Rikers' et al. (1997) study it was predicted that expert physicians cannot diagnose a case outside their specialty in an encapsulating mode, and therefore they would perform at an intermediate level of expertise. This "relapse" of expert physicians should be expressed in a performance that is not different from that of advanced students. Surprisingly, this study showed no significant differences in the applied encapsulating concepts between experts and subexperts, albeit significant differences in diagnostic accuracy and processing speed were found (cf. Van de Wiel, Boshuizen, & Schmidt, 1998). This finding suggest that experts and subexperts alike process clinical cases in essentially the same way. However, the present study shows that this view is not the entirely justified. Response times and error rates indicate that the expert physician's medical knowledge (i.e., biomedical and encapsulated knowledge) is better organized than that of subexperts. It seems, therefore, that the "traditional" measures of free recall and pathophysiology are not sensitive enough to detect the often small differences between experts and subexperts. That is, both experts and subex-

perts posses encapsulated knowledge, but expert physicians are able to activate it faster.

Another viable explanation why previous clinical case studies comparing experts and subexperts did not reveal significant differences is provided by Patel and colleagues (1990, 1991). They assume that experts and subexperts alike, have developed some kind of generic expertise that enables them to process and retain relevant case information. Generic expertise implies the existence of a schema or script that is based on their practical experience. Therefore, subexperts are able to recall and explain the sign and symptoms of a clinical case just as well as do expert physicians. Again, the present study shows that this view is only partly true. Response times and error rates indicate that the clinical case representation of expert physicians is better organized than that of subexperts and students. Therefore, it is more likely that the traditional tasks are too generic than that the physicians' knowledge is too generic.

In the discussion section of Experiment 1 a distinction was made between on-line and off-line high-level inferences (Graesser & Bower, 1990; Graesser, Singer, & Trabasso, 1994; McKoon & Ratcliff, 1992). It was assumed that the expert physicians made their high-level inferences during comprehending the clinical case, subexperts and medical students, on the other hand, made these inferences off-line (i.e., during the later decision task). The significant difference in response times seems to support this assumption. However, based on the methodology used in the experiments we cannot exclude the possibility that also expert physicians made some of their inferences off-line. This issue is especially pertinent because, if expert physicians also made most of their inferences off-line, their clinical case representation cannot be characterized based on encapsulated knowledge. Further research avoiding the possibility of off-line inferences is therefore necessary in order to clarify this issue.

In sum, findings of the present study indicate that the main difference between experts, subexperts and medical students, is not the presence or absence of encapsulated knowledge, but the speed and accuracy of the availability of encapsulated knowledge. Further research is necessary to answer the question whether the activation of encapsulated knowledge of experts, subexperts and students takes place during case comprehension or during the later decision task.

References

- Blessing, S. B., & Anderson, J. R. (1996). How do people learn to skip steps? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22 (3), 576-598.
- Bordage, G. (1994). Elaborated knowledge: A key to successful diagnostic thinking. *Academic Medicine*, 69(11), 883-885.
- Boshuizen, H. P. A. (1989). *De ontwikkeling van medische expertise; een cognitief-psychologische benadering*. [On the development of medical expertise; a cognitive psychological approach]. Doctoral dissertation. Haarlem: Thesis.
- Boshuizen, H. P. A., & Schmidt, H. G. (1992). On the role of biomedical knowledge in clinical reasoning by experts, intermediates and novices. *Cognitive Science*, 16, 153-184.
- Claessen, H. F. A., & Boshuizen, H. P. A. (1985). Recall of medical information by students and doctors. *Medical Education*, 19, 61-67.
- Collins, A. M., & Loftus, E. R. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407-428.
- Coughlin, L. D., & Patel, V. L. (1986, April). *Text comprehension and expertise in the domain of medicine*. Paper presented at the Meeting of the American Educational Research Association, San Francisco, CA.
- Custers, E. J. F. M., Boshuizen, H. P. A., & Schmidt, H. G. (1996). The influence of medical expertise, case typicality, and illness script component on case processing and disease probability estimates. *Memory and Cognition*, 24(3), 384-399.
- Elstein, A.S., Schulman, L.S., & Sprafka, S.A. (1978). *Medical problem solving: An analysis of clinical reasoning*. Cambridge, MA: Harvard University Press.
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102(2), 211-245.
- Feltovich, P.J. & Barrows, H.S. (1984). Issues of generality in medical problem solving. In: Schmidt, H.G. & De Volder, M.L. (Eds.), *Tutorials in problem-based learning* (pp. 128-142). Assen: Van Gorcum.
- Graesser, A. C., & Bower, G.H. (Eds.). (1990). *Inferences and text comprehension*. San Diego, CA: Academic Press.
- Graesser, A.C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101, 371-395.
- Koedinger, K. R., & Anderson, J. R. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science*, 14, 511-550.
- Lesgold, A.M. (1984). Acquiring expertise. In J.R. Anderson & S.M. Kosslyn (Eds.), *Tutorials in learning and memory: Essays in honor of Gordon Bower*, (pp 31-60). San Francisco, CA: Freeman.
- Lesgold, A.M., Rubinson, H., Feltovich, P.J., Glaser, R., Klopfer, D. & Wang, Y. (1988). Expertise in a complex skill: Diagnosing X-ray pictures. In M. T. H. Chi, R. Glaser, & M. Farr (Eds.), *The nature of expertise*, (pp. 311-342). Hillsdale, NJ: Erlbaum.
- McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, 99, 440-446.

- Norman, G. R., Trott, A. D., Brooks, L. R., & Smith, E. K. M. (1994). Cognitive differences in clinical reasoning related to postgraduate training. *Teaching and Learning in Medicine*, 6(2), 114-120.
- Patel, V. L., Arocha, J. F., & Groen, G. J. (1986). Strategy selection and degree of expertise in medical reasoning. In C. Clifton (Ed.), *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*, (pp. 780-791). Hillsdale, NJ: Erlbaum.
- Patel, V. L., Evans, D. A., and Chawla, A. S. (1986). Predictive versus diagnostic reasoning in the application of biomedical knowledge. In: *Proceedings of the 8th Annual Conference of the Cognitive Science Society*, Lawrence Erlbaum, Hillsdale, NJ.
- Patel, V. L., Evans, D. A., & Groen, G. J. (1989). Biomedical knowledge and clinical reasoning. In D. A. Evans & V. L. Patel (Eds.), *Cognitive science in medicine: Biomedical modeling*, (pp. 53-112). Cambridge, MA: MIT Press.
- Patel, V. L., & Groen, G. J. (1986). Knowledge based solution strategies in medical reasoning. *Cognitive Science*, 10, 91-116.
- Patel, V. L., & Groen, G. J. (1991). Developmental accounts of the transition from medical student to doctor: Some problems and suggestions. *Medical Education*, 25(6), 527-535
- Patel, V. L., Groen, G. J., & Arocha, J. F. (1990). Medical expertise as a function of task difficulty. *Memory and Cognition*, 18(4), 394-406.
- Patel, V. L., & Kaufman, D. R. (1995). Clinical reasoning and biomedical knowledge: Implications for teaching. In J. Higgs & M. Jones (Eds.), *Clinical reasoning in the health professions*, (pp. 117-128). Oxford, UK: Butterworth Heinemann.
- Patel, V. L., & Medley-Mark, V. (1986). Relationship between representation of textual information and underlying problem representation in medicine. *Cognitive research series*. Montreal: McGill University, Centre for Medical Education. (CME86-CS1).
- Rikers, R. M. J. P., Boshuizen, H. P. A., & Schmidt, H. G. (1997, March). *Nonroutine problem solving by medical experts*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, ILL.
- Rikers, R. M. J. P., Schmidt, H. G., & Boshuizen, H. P. A. (in press). Knowledge encapsulation and the intermediate effect. *Contemporary Educational Psychology*.
- Schmidt, H. G., & Boshuizen, H. P. A. (1992). Encapsulation of biomedical knowledge. In D. A. Evans & V. L. Patel (Eds.), *Advanced models of cognition for medical training and practice*. New York, NY: Springer Verlag.
- Schmidt, H. G., & Boshuizen, H. P. A. (1993). On the origin of intermediate effects in clinical case recall. *Memory and Cognition*, 21(3), 338-351.
- Schmidt, H. G., Boshuizen, H. P. A., & Norman, G. R. (1992). Reflections on the nature of expertise in medicine. In E. Keravnou (Ed.), *Deep models for medical knowledge engineering*. Amsterdam, The Netherlands: Elsevier.
- Van de Wiel, M. W. J. (1997). Knowledge encapsulation: Studies on the development of medical expertise. Doctoral dissertation. Wageningen: Ponsen & Looijen.
- Van de Wiel, M. W. J., Schmidt, H. G., & Boshuizen, H. P. A. (1998). A failure to reproduce the intermediate effect in clinical case recall. *Academic Medicine*, 73(8), 894-900.

Van de Wiel, M.W.J., Boshuizen, H.P.A., Schmidt, H.G., & de Leeuw, P.W. (1993).
A failure to reproduce the intermediate effect in clinical case recall. Paper presented
at the Meeting of the American Educational Research Association, Washington,
DC.

Appendix 5.1

Case of Acute Bacterial Endocarditis

<i>Signs & Symptoms</i>	<i>Encapsulating Items</i>
Diastolic murmur	Endocarditis
High pulse rate	Fever reaction
Loss of vision	Aortic insufficiency
Normal CVP	Drug abuse
Puncture wounds	Acute & Severe
Shaking chills	Sepsis
Sweating	Thromboembolism
Temperature 40° C (104° F)	Valvular vegetation

Case of Heart Failure

<i>Signs & Symptoms</i>	<i>Encapsulating Items</i>
CPK normal	Backward failure
Enlarged left atrium	Cardiomyopathy
Enlarged left ventricle	Forward failure
High hemoglobin level	Heart failure
Holosystolic murmur	Left-sided heart failure
Hypotension	Mitral valve insufficiency
Low ESR level	Right-sided heart failure
Rales	Venous congestion

Appendix 5.2

Items related to clinical text of congestive heart failure

<i>Biomedical Items</i>	<i>Encapsulating Items</i>
Myocardial ischemia	Systolic dysfunction
Vasoconstriction	Diastolic dysfunction
Salt retention	Circulatory failure
Water retention	Neurohumoral changes
↑ contractility myocyte	Coronary sclerosis
Sympathicotonia	Cardiac Hypertrophy
↑ diastolic pressure	Mitral valve insufficiency
↑ wall stress	Aortic insufficiency
↓ contractility	High-output failure
↑ filling pressure	Aortic stenosis
↑ stroke volume	Mitral stenosis
↑ renin	Cor pulmonale
↑ ADH	Cardiomyopathy
↑ preload	Constrictive pericarditis
Insufficient oxygen supply	External pericarditis
Tissue degeneration	Left-sided heart failure
↑ angiotensin	Right-sided heart failure
↓ O ₂ perfusion	Heart failure
Left-ventricle dilatation	Forward failure
Right ventricle dilatation	Backward failure

Appendix 5.3

Clinical Text of Congestive Heart Failure

Heart failure occurs when the heart loses its ability to pump enough blood through the body. Usually, the loss in pumping action is a symptom of an underlying heart problem, such as coronary artery disease. The term heart failure suggests a sudden and complete stop of heart activity. But, actually, the heart does not suddenly stop. Rather, heart failure usually develops slowly, often over years, as the heart gradually loses its pumping ability and works less efficiently. Some people may not become aware of their condition until symptoms appear years after their heart began its decline. How serious the condition is depends on how much pumping capacity the heart has lost. Nearly everyone loses some pumping capacity as they age. But the loss is significantly more in heart failure and often results from a heart attack or other disease that damages the heart. The severity of the condition determines the impact it has on a person's life. At one end of the spectrum, the mild form of heart failure may have little effect on a person's life; at the other end, severe heart failure can interfere with even simple activities and prove fatal. Between those extremes, treatment often helps people lead full lives. But all forms of heart failure, even the mildest, are a serious health problem, which must be treated. To improve their chance of living longer, patients must take care of themselves, see their physician regularly, and closely follow treatments.

The term congestive heart failure is often used to describe all patients with heart failure. In reality, congestion (the buildup of fluid) is just one feature of the condition and does not occur in all patients. There are two main categories of heart failure--systolic and diastolic. However, within each category, symptoms and effects may differ from patient to patient. The two categories are systolic heart failure and diastolic heart failure. A number of symptoms are associated with heart failure, but none is specific for the condition. Perhaps the best known symptom is shortness of breath ("dyspnea"). In heart failure, this may result from excess fluid in the lungs. The breathing difficulties may occur at rest or during exercise. In some cases, congestion may be severe enough to prevent or interrupt sleep. Fatigue or easy tiring is another common symptom. As the heart's pumping capacity decreases, muscles and other tissues receive less oxygen and nutrition, which are carried in the blood. Without proper "fuel," the body cannot perform as much work, which translates into fatigue.

Fluid accumulation, or edema, may cause swelling of the feet, ankles, legs, and occasionally, the abdomen. Excess fluid retained by the body may result in weight gain, which sometimes occurs fairly quickly. Persistent coughing is another common sign, especially coughing that regularly produces mucus or pink, blood-tinged sputum. Some people develop raspy breathing or wheezing. Because heart failure usually develops slowly, the symptoms may not appear until the condition has progressed over years. The heart hides the underlying problem by making adjustments that delay--but do not prevent--the eventual loss in pumping capacity. The heart adjusts, or compensates, in different ways to cope with and hide the effects of heart failure:

By making these adjustments, or compensating, the heart can temporarily make up for losses in pumping ability, sometimes for years. However, compensation has its limits. Eventually, the heart cannot offset the lost ability to pump blood, and the signs of heart failure appear. (*Source: U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES, Public Health Service, National Institutes of Health, National Heart, Lung, and Blood Institute, NIH Publication No. 94-923, April 1994*).

Chapter 6

Discussion and Conclusions

This thesis concerns the role of so-called "encapsulated knowledge" within and outside the expert physicians' domain of expertise. Previous clinical case studies have demonstrated that in order to become an expert physician important changes occur in the students' medical knowledge (Schmidt & Boshuizen, 1992, 1993; Schmidt, Boshuizen, & Norman, 1992; Van de Wiel, Boshuizen, & Schmidt, 1998). According to Schmidt and Boshuizen (1992, 1993), medical students undergo a stage-like transition from the first day they enter medical school until they have reached the level of expert physician. During their first years in medical school students acquire through lectures and individual study primarily biomedical knowledge (i.e., knowledge related to the processes underlying the manifestations of disease). Therefore, their diagnostic reasoning consists of long chains of small steps based on detailed biomedical concepts. Their clinical knowledge (i.e., knowledge related to the manifestations of disease), however, is limited to descriptions in textbooks, because they have not encountered real patients. Expert physicians, on the other hand, have extensive clinical knowledge. Their biomedical knowledge has, as a result of the confrontation with real patients, become subsumed under a small number of clinically relevant concepts, the so-called encapsulating concepts. That is, elements of their biomedical knowledge have become interconnected so that it can be accessed as coherent chunks of information, rather than fragments. By the time students have reached the level of an expert physician, it is not necessary anymore to activate all possible relevant knowledge in order to understand the signs and symptoms, only knowledge that is pertinent to the case will be activated (cf. Blessing & Anderson, 1996; Koedinger & Anderson, 1990). Therefore, expert physicians that have to diagnose a clinical case will mainly activate specific pre-existing encapsulating concepts (Figure 6.1.).

Evidence to support the encapsulation of the physician's knowledge structure has been found in free recall, pathophysiological explanations, and think-aloud studies. Most eminently in this respect is the so-called "intermediate

effect" (e.g., Boshuizen, 1989; Coughlin & Patel, 1986; Patel & Medley-Mark, 1986; Schmidt & Boshuizen, 1993; Schmidt, Boshuizen, & Hobus, 1988).

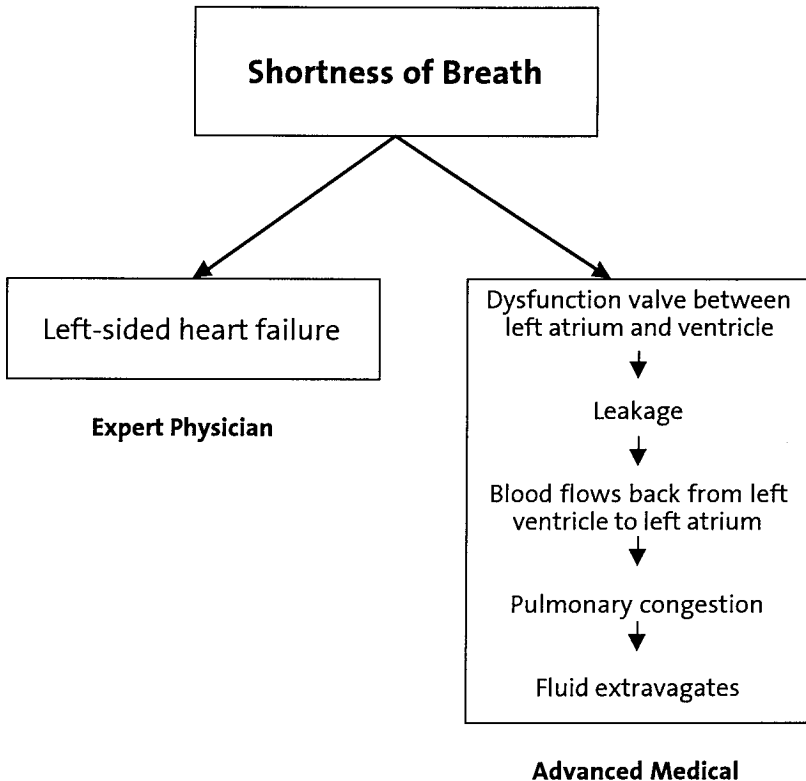


Figure 6.1. In order to account for the patient's shortness of breath, the advanced medical students has to engage in biomedical reasoning. The expert physician, on the other hand, can skip the student's detailed steps by using an encapsulating concept (Van de Wiel, 1997).

If participants of different levels of expertise are instructed to write down everything they remember or explain the signs and symptoms in a clinical case, intermediates (i.e., advanced students) recall more and provide more elaborate pathophysiological explanations than do expert physicians. This is a rather counter-intuitive finding, because most theories on the development of expertise predict a linear performance with level of expertise (e.g., Chase & Simon, 1973; Ericsson & Kintsch, 1995; Ericsson & Staszewski, 1989). However, this finding can easily be explained by the idea of knowledge encapsulation. Expert physicians, in contrast to students, apply many

encapsulating concepts in recall and pathophysiology, thereby summarizing the essential information in the case.

Although the role of encapsulated knowledge within the physicians' domain of expertise has been studied extensively, its role outside the physicians' domain of expertise remains unclear. In order to clarify this issue, the present thesis reports four studies on the role of encapsulated knowledge outside the borders of the expert physician's domain of expertise. Based on the theory of knowledge encapsulation it was predicted that experts process a clinical case within their specialty in an encapsulated mode. However, outside their specialty they will relapse into an elaborate biomedical processing approach, which is also typical for advanced students.

In Chapter 2 expert physicians from two specialties (i.e., cardiologists and pulmonologists) had to study four clinical cases, provide a diagnosis, recall everything they remembered, and finally explain the signs and symptoms in the case. Two cases were within the domain of cardiology and two were pulmonary cases. Therefore, cardiologists and pulmonologists alike fulfilled both the role of experts and that of subexperts in this study. Contrary to our expectation, findings were not completely in line with knowledge encapsulation theory. Although experts were faster and more accurate in their diagnoses than subexperts, recall and pathophysiological explanations did not reveal significant differences. Findings were explained by the complexity of the clinical cases: Experts were not able to process the cases in a routine manner (i.e., encapsulated mode) and, therefore, diagnosed the cases in an elaborate biomedical mode, in a similar manner as did the subexperts. An alternative explanation that was also discussed, is provided by Patel and colleagues (Patel & Groen, 1991a, 1991b; Patel, Groen, & Arocha, 1990). They assume that experts and subexperts alike have developed so-called "generic expertise." That is, experts and subexperts have developed schema's belonging to a particular disease, which enable them to retain the essential information in a clinical case, within or outside their specialty. In contrast to subexperts, domain experts also have "specific expertise," which enables them to link the non-salient cues or loose ends to the main diagnosis, and thereby providing a more accurate diagnosis. A third explanation that was discussed might be the lack of "remoteness" between the two specialties of cardiology and pulmonology. This is especially true in the Netherlands, where cardiologists and pulmonologists share the first two years of their residency training. This is a period where they both have to evaluate cardiological and pulmonological problems on a regular basis. Possibly, encapsulation starts quite early in their residency, resulting in equal levels of encapsulated knowledge for both specialties.

The study reported in Chapter 3 tried to remedy the problem related to the

difficulty of the clinical cases and that of the remoteness of the specialties. Two clinical cases were used that had been used in previous clinical studies (e.g., Patel and Groen, 1986; Schmidt and Boshuizen, 1993; Van de Wiel, Boshuizen, & Schmidt, 1998). Again, expert physicians from two different specialties were instructed to study, recall and explain two clinical cases. In contrast to Experiment 1 only one group fulfilled the role of experts (cardiologists) the other group were the subexperts (neurologists). The physicians had three minutes to study each clinical case. In contrast to the previous study, findings were in line with encapsulating theory. Experts were more accurate in their diagnoses than were subexperts. More importantly, experts introduced more encapsulating concepts in their recall and pathophysiological explanations.

However, although these findings can be explained fairly straightforward with encapsulation theory, some critical remarks seem justified. First, experts diagnostic accuracy was not very high (about 73 % of the maximum score), and only two cardiologists provided a completely correct diagnosis on both cases. This indicates that the cases were difficult problems for the experts in this study and it is, therefore, not unlikely that experts were not completely able to process these clinical cases in a routine manner. Second, the number of encapsulating concepts in the expert physician's recall is low (about 17 % of the total recall), albeit significantly higher than that of subexperts. This cast some doubt on the contribution of encapsulating concepts to the differences in the total number of propositions recalled by experts and subexperts. So, if the proportion of encapsulating concepts is taken into consideration, it should be concluded that experts' recall is not "dominated" by encapsulating concepts. Finally, the number of concepts and links within the experts' pathophysiological explanation does not significantly differ from that of subexperts. Notwithstanding, experts did roughly introduce twice as many encapsulating concepts in their explanation than did subexperts. Again, this seems to indicate that the elaborateness of the physicians' explanations seems relatively independent of the number of encapsulating concepts. This issue is especially pertinent, because the *intermediate effect* as it was originally found, was solely attributed to the influence of encapsulated knowledge.

The study reported in Chapter 4 investigates the relation between encapsulated knowledge and the intermediate effect. If encapsulated knowledge is confined to the physicians' domain of expertise, the intermediate effect should not appear outside the physicians' domain of expertise. That is, experts outside their specialty are expected to perform very much like advanced medical students. Second-year, sixth-year medical students, and neurologists were instructed to diagnose, recall, and explain four clinical case

descriptions outside the neurologists' domain of expertise. However, findings seem to indicate that experts are relatively insensitive for the nature of the cases used: An intermediate effect was found for the recall data and pathophysiological explanations. The measure of diagnostic accuracy; on the other hand, showed a linear relation with expertise level. This study sheds a new light on the intermediate effect, because this phenomenon seems to occur within and outside the physician's domain of expertise. Furthermore, if encapsulated knowledge is considered as the only constituent of the intermediate effect, this study and those described in Chapters 2 and 3, cast some serious doubt on the role of encapsulated knowledge to the intermediate effect.

The question, then, is what are the constituents of the intermediate effect if knowledge encapsulation is not the only cause? A possible answer to this question might be found in the differences in goals and motivation of expert physicians and medical students. Schmidt (1982, 1984), for instance, proposed a *cognitive set-point hypothesis*. This hypothesis assumes that during information processing a cognitive set-point is applied, indicating when sufficient study time has been spent on processing certain information in order to recall it at a later stage. The cognitive set-point is reached when a certain amount of time has been spent on processing the information. The set-point is supposed to be determined by an estimation of some minimum amount of understanding by the readers, their goals and motivation, and external pressures such as the expected consequences of not being able to reproduce adequately the information. If the set-point level for a particular information unit has been reached, the reader will proceed to process new information until the set-point for all information units has been reached or time has run out. Therefore, if participants have high levels of prior knowledge they would require less time to reach their set-point for the information (Machiels-Bongaerts, 1993).

In order to investigate the psychological plausibility of encapsulated knowledge, and to overcome the problems associated with the "traditional" clinical case studies, a more direct approach was undertaken to probe the physicians' clinical case representation. Chapter 5 describes two priming experiments that deviate from the traditional research paradigm consisting of diagnosis, free recall and pathophysiological explanation. That is, instead of recalling or explaining the clinical case description, participants were instructed to decide as fast as possible whether a medical concept presented on a computer screen is related to the clinical case. Within this framework the clinical case description acts as "prime" for the knowledge participants activate during case processing. The resulting pre-activation of their medical knowledge is assumed to facilitate the subsequent evaluation process that requires access to this information in memory, thereby producing a *semantic*

priming effect (Collins & Loftus, 1975). If the encapsulation theory is correct in assuming that expert physicians build a clinical case representation based on pre-existing encapsulating concepts within their specialty, experts are expected to make faster and more accurate responses while evaluating case-related encapsulating items, than do subexperts and medical students. After all, the latter two groups have to resort to their biomedical knowledge in order to build an adequate clinical case representation (Schmidt & Boshuizen, 1992, 1993). In the first priming experiment, consisting of two clinical cases, participants had one minute to study each clinical case. One half of the participants got the correct diagnoses of the cases in advance (i.e., stated condition), the other half had to provide the diagnoses themselves (i.e., unstated condition). Subsequently, all participants had to evaluate three types of items on their relatedness to a clinical case: Literally stated signs and symptoms, inferred encapsulating items (i.e., were no part of the case), and filler items, which were not related to the clinical case. The outcome of this study was in line with encapsulation theory: Experts were indeed faster and more accurate than students and subexperts in evaluating encapsulating items. Interestingly, experts were also faster and more accurate in evaluating literally stated signs and symptoms. The opposite relation was true for the subexperts: They evaluated the signs and symptoms faster than the encapsulating items. Students, on the other hand, did not show a "preference" for either type of items.

Surprisingly, the diagnostic accuracy of the expert physicians in the diagnostic unstated condition was not very high (about 50 % of maximum score), albeit significantly higher than both other groups. This might be an indication that expert physicians, who did not get the correct diagnoses in advance (i.e., the unstated condition), did not entirely process these cases in an encapsulated mode. Otherwise, differences in response times between informed and uninformed experts would not have been so likely. Furthermore, the impact of the presence of the correct diagnoses seems to equally aid all participants. Students, subexperts, and experts were about 400 ms faster on both types of items if they were informed about the correct diagnosis than when they were not. Therefore, stating the correct diagnosis seems to equally prime the encapsulating items, and signs and symptoms.

However, although the first priming experiment provided more insight into the role of encapsulated knowledge and signs and symptoms, the role of biomedical knowledge remains unclear within the physicians' clinical case representation.

The second experiment of Chapter 5, therefore, investigated the role of biomedical knowledge and encapsulated knowledge within the physicians' knowledge structure. Schmidt and Boshuizen (1992, 1993) assumed that

physicians relapse into an elaborate biomedical processing approach if they are confronted with problems outside their specialty. Fourth-year medical students, subexperts, and expert physicians were instructed to study a text describing signs and symptoms relating to the disease of heart failure. According to Schmidt and Boshuizen's model of knowledge encapsulation, signs and symptoms in a case activate specific pre-existing encapsulating concepts in the expert's physicians LTM. Students and Subexperts, on the other hand, who cannot rely on encapsulated knowledge, have to resort to their biomedical knowledge in order to understand the signs and symptoms. Again, three types of items were used: Encapsulating items, biomedical items, and filler items. Both encapsulating items and biomedical items were not explicitly stated in the text. The results showed that expert physicians were faster than both other groups in evaluating encapsulating and biomedical items. Interestingly, medical students and subexperts, although considerably slower, showed the same tendency in their response times as did the expert physicians. They also evaluated encapsulating items faster than biomedical items. Furthermore, there were no significant differences in the mean error rates between the different levels of expertise.

In conclusion, findings of both priming studies provide us with new a perspective on the studies using the traditional clinical case paradigm. First of all, priming studies seem to suggest that both experts and subexperts and even medical students possess encapsulated knowledge, and thereby processing a clinical case in essentially the same way. Encapsulated knowledge is, therefore, no unique characteristic of expert physicians' knowledge structure. However, taken the response times into consideration, we also have to conclude that encapsulating concepts do have a special position within the expert physicians' knowledge structure. Second, the priming approach seems a more promising approach to detect differences in clinical case representations between experts and subexperts. As has been outlined before, previous clinical case studies based on the traditional measures, showed inconclusive results (see Chapter 2). That is, measures of free recall and pathophysiology might not capture all (encapsulating) concepts that experts, subexperts, and students activate during case processing. However, this is to a certain extent also true for the current priming approach. The present procedure of a case followed by different types of items, makes it difficult to distinguish between 'on-line' inferences (i.e., inferences that are made during the course of case comprehension) and 'off-line' inferences (i.e., inferences that are made during the later decision task). If we take a closer look at the duration of the response times it is not unlikely that experts and subexperts alike deduced the inferences at the moment they were confronted with the items, and, hence, these items were no part of the initial case representation. Possibly, the only differ-

ence between experts and subexperts is the speed of inferring the validity of encapsulating items.

Conclusions

A first conclusion that could be drawn from the findings of the present studies is that they support the theory of knowledge encapsulation to a large extent. Especially the findings of the priming studies support the unique position of encapsulating concepts within the expert physicians' knowledge structure. Experts are faster and more accurate in evaluating encapsulating concepts than subexperts and medical students. However, encapsulating concepts are not a characteristic unique to the expert physicians' representation of a clinical case. Both subexperts and medical students also possess and apply encapsulated knowledge in their clinical case representations to some extent. The important difference between experts and subexperts and students is, therefore, not whether or not they possess encapsulating concepts, but the extent to which encapsulating concepts have been activated while processing a clinical case. Findings suggest that experts' encapsulating concepts have reached higher levels of activation than those of subexperts and medical students have.

A second conclusion that could be drawn from the present studies is that measures such as processing speed and diagnostic accuracy are reliable and stable measures to detect differences between different levels of expertise. A finding that is also corroborated by previous clinical case studies (e.g., Claessen & Boshuizen, 1985; Schmidt & Boshuizen, 1993; Van de Wiel, 1997). The measures of free recall and pathophysiology, on the other hand, are more problematic. For instance, the study described in Chapter 2 did not show significant differences between expert and subexperts, although they were predicted. However, Chapter 3 describes a study between experts and subexperts that were in line with the predications derived from encapsulation theory. Taken these different outcomes into consideration, it should be concluded that recall and pathophysiology are not entirely reliable measures to investigate differences between experts and subexperts. Priming studies such as described in Chapter 5 seem therefore more appropriate to this end. A third conclusion that could be drawn is that the relapse of expert physicians outside their specialty cannot be solely attributed to the failure to activate encapsulating concepts. First of all, because the study described in Chapter 2 failed to detect differences between experts and subexperts, indicating that expert physicians outside their domain of expertise do not relapse at all. Secondly, the study described in Chapter 4 demonstrated, for the first

time, an intermediate effect outside the expert physicians' specialty. Alternative explanations such as provided by Patel and Groen (1991a) seem therefore more promising: Experts filter out irrelevant information during case processing. Medical students, on the other hand, report also information that is of little diagnostic relevance.

Suggestions for further research

The research presented in this thesis resulted in a number of suggestions for further research.

The present studies showed that the traditional clinical case approach of diagnosis, recall, and pathophysiological explanations might result in divergent findings concerning the role of encapsulated knowledge. The study described in Chapter 2, for instance, revealed no significant differences between experts and subexperts concerning the application of encapsulated knowledge. The study reported in Chapter 3, on other hand, was considerably more in line with encapsulation theory. As has been outlined before, these divergent findings might be explained by the lack of sensitivity of the post-hoc measures of free recall and pathophysiological explanations. A measure that was not used in the present studies, but has proven to be successful in previous clinical case studies comparing medical students and expert physicians, is think-aloud (i.e., while diagnosing a clinical case, participants are instructed to report everything that comes to their mind). Verbal protocols obtained by thinking aloud are assumed to reveal information that is available during comprehension of the clinical case (Ericsson, 1988; Ericsson & Simon, 1984; Fletcher, 1986). The think-aloud methodology might prevent experts and subexperts from using a kind of "output editing." That is, thinking aloud might provide us with a tool that enables us to gain more insight of the build-up of a clinical case representation, minimizing the opportunity for experts and subexperts to withhold or edit information that is activated while they process the clinical case.

Furthermore, the study time of a clinical case might also be a good way to manipulate the clinical case processing. The study described in Chapter 2 did not reveal significant differences between experts and subexperts in recall and pathophysiology, but it did show clear and significant differences in study time: Subexperts needed considerably more time to diagnose a clinical case than did expert physicians. Therefore, if subexperts get less time to process a clinical case, the differences in case representation might be revealed (cf. Schmidt & Boshuizen, 1993). That is, subexperts are able to build-up an encapsulated clinical case representation only if they have enough time to do

so. Evidence to support this hypothesis is found in the study described in Chapter 3, where study time was restricted.

The priming approach, however, has proven to be the most fruitful approach in clinical case studies to investigate clinical case representations by experts and subexperts. It provides the opportunity to probe the physician's clinical case description more directly than is possible with the traditional clinical case approach. However, the study described in Chapter 5 has one important shortcoming: It is not possible to make a distinction between inferences made on-line or off-line. That is, it is not possible to decide whether the inference was made during the original processing of the case or during the later decision task (Graesser & Bower, 1990; Graesser, Singer, & Trabasso, 1994; McKoon & Ratcliff, 1992; Suh & Trabasso, 1993). In order to remedy this problem, the physicians representation should be probed during the case processing. So, instead of studying a complete clinical case description followed by a decision task, clinical cases should be subdivided into smaller parts that are presented to the physicians. Subsequently, after reading such a short passage of a clinical case, possible inferences based on the passage should be presented on a screen for evaluation. This approach might also overcome the problem that the activation level of relevant concepts decays, because the interval between the initial activation and the later decision task is too large. Furthermore, the priming approach might provide us with a tool to investigate the relationship between biomedical knowledge and clinical knowledge more deeply. Previous studies that investigated the relationship between biomedical and clinical knowledge did not reveal conclusive evidence (e.g., Lesgold, 1984; Lesgold, Rubinson, Feltovich, Glaser, Klopfer, & Wang, 1988; Patel, Evans, & Chawla, 1986; Patel & Groen, 1986; Patel, Groen, & Arocha, 1990). That is, the status of biomedical knowledge and clinical knowledge, and their mutual relation, is still unclear within the physicians' clinical case representation. In this respect, the priming approach might provide us with a useful tool to investigate the physician's mind more deeply.

A final suggestion for further research might be the application of encapsulation theory beyond the borders of medical expertise (cf. Van de Wiel, 1997). Although the theory of knowledge encapsulation has its roots within the medical domain, this does not necessarily mean that its value is restricted to this domain. For instance, Abelson (1984) has shown that expert computer programmers do not store details of the code. She found that novices actually had better memory for details of code than did experts. The reason appeared to be that experts attended more to the overall goal structure of the programming task, rather than the actual code. This finding indicates that experts only have to pay attention to the overall goal structure of the program in order to understand its functioning. Hence, experts do not have to

engage in detailed analysis of the code in order to understand the program.¹ Similarly, in the medical domain, where experts do not have to engage in detailed biomedical reasoning in order to understand the patient's disease.

In conclusion, the expert-subexpert approach to investigate the physicians' mental representations might turn out be fruitful, because it stimulates the development of more advanced and sensitive research methods, which might enables us to gain an unique insight of the prospects and limitations of medical expertise.

References

- Adelson, B. (1984). When novices surpass experts: The difficulty of a task may increase with expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 483-495.
- Blessing, S. B., & Anderson, J. R. (1996). How do people learn to skip steps? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22 (3), 576-598.
- Boshuizen, H. P. A. (1989). *De ontwikkeling van medische expertise; een cognitief-psychologische benadering*. [On the development of medical expertise; a cognitive psychological approach]. Doctoral dissertation. Haarlem: Thesis.
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55-81.
- Claessen, H. F. A., & Boshuizen, H. P. A. (1985). Recall of medical information by students and doctors. *Medical Education*, 19, 61-67.
- Collins, A. M., & Loftus, E. R. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407-428.
- Coughlin, L. D., & Patel, V. L. (1986, April). *Text comprehension and expertise in the domain of medicine*. Paper presented at the Meeting of the American Educational Research Association, San Francisco, CA.
- Ericsson, K. A. (1988). Concurrent verbal reports on text comprehension: A review. *Text*, 8, 295-235.
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102(2), 211-245.
- Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: verbal reports as data*. Cambridge, MA: MIT Press.
- Ericsson, K.A., & Staszewski, J. (1989). Skilled memory and expertise: Mechanisms of exceptional performance. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp. 235-267). Hillsdale, NJ: Erlbaum.

¹ In fact, encapsulation, although not completely comparable with its use in clinical case studies, is an important feature of object-oriented programming languages. Encapsulation, sometimes referred to as data hiding, is the mechanism whereby the implementation details of some part of the code are kept hidden from the user.

- Graesser, A. C., & Bower, G.H. (Eds.). (1990). Inferences and text comprehension. San Diego, CA: Academic Press.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994) Constructing inferences during narrative text comprehension. *Psychological Review*, 101(3), 371-395.
- Fletcher, C. R. (1986). Strategies for the allocation of short-term memory during comprehension. *Journal of Memory and Language*, 27, 235-244.
- Koedinger, K. R., & Anderson, J. R. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science*, 14, 511-550.
- Lesgold, A.M. (1984). Acquiring expertise. In J.R. Anderson & S.M. Kosslyn (Eds.), *Tutorials in learning and memory: Essays in honor of Gordon Bower*, (pp 31-60). San Francisco, CA: Freeman.
- Lesgold, A.M., Rubinson, H., Feltovich, P.J., Glaser, R., Klopfer, D. & Wang, Y. (1988). Expertise in a complex skill: Diagnosing X-ray pictures. In M. T. H. Chi, R. Glaser, & M. Farr (Eds.), *The nature of expertise*, (pp. 311-342). Hillsdale, NJ: Erlbaum.
- Machiels-Bongaerts, M. (1993). Mobilizing prior knowledge in text processing: The selective-attention hypothesis versus the cognitive set-point hypothesis. University Press Maastricht.
- McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, 99, 440-446.
- Patel, V. L., Evans, D. A., and Chawla, A. S. (1986). Predictive versus diagnostic reasoning in the application of biomedical knowledge. In: *Proceedings of the 8th Annual Conference of the Cognitive Science Society*, Lawrence Erlbaum, Hillsdale, NJ.
- Patel, V. L., & Groen, G. J. (1986). Knowledge based solution strategies in medical reasoning. *Cognitive Science*, 10, 91-116.
- Patel, V. L., & Groen, G. J. (1991a). The general and specific nature of medical expertise: A critical look. In A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits*, (pp. 93-125). New York: Cambridge University Press.
- Patel, V. L., & Groen, G. J. (1991b). Developmental accounts of the transition from medical student to doctor: Some problems and suggestions. *Medical Education*, 25(6), 527-535.
- Patel, V. L., Groen, G. J., & Arocha, J. F. (1990). Medical expertise as a function of task difficulty. *Memory and Cognition*, 18(4), 394-406.
- Patel, V.L., & Medley-Mark, V. (1986). Relationship between representation of textual information and underlying problem representation in medicine. *Cognitive research series*. Montreal: McGill University, Centre for Medical Education. (CME86-CS1).
- Schmidt, H. G. (1982). Activation and restructuring of prior knowledge and their effects on text processing. In A. Flammer, & W. Kintsch (Eds.). *Discourse Processing*. Amsterdam: North-Holland.

- Schmidt, H. G. (1984). Activatie van voorkennis en tekstverwerking [Prior knowledge activation and text processing]. *Nederlands tijdschrift voor de psychologie*, 39, 335-347.
- Schmidt, H. G., & Boshuizen, H. P. A. (1992). Encapsulation of biomedical knowledge. In D. A. Evans & V. L. Patel (Eds.), *Advanced models of cognition for medical training and practice*. New York, NY: Springer Verlag.
- Schmidt, H. G., & Boshuizen, H. P. A. (1993). On the origin of intermediate effects in clinical case recall. *Memory and Cognition*, 21(3), 338-351.
- Schmidt, H. G., Boshuizen, H. P. A., & Hobus, P. P. M. (1988). Transitory stages in the development of medical expertise: The "intermediate effect" in clinical case representation studies. In: *Proceedings of the Cognitive Science Society Meeting*, Lawrence Erlbaum, Hillsdale, NJ.
- Schmidt, H. G., Boshuizen, H. P. A., & Norman, G. R. (1992). Reflections on the nature of expertise in medicine. In E. Keravnou (Ed.), *Deep models for medical knowledge engineering*. Amsterdam, The Netherlands: Elsevier.
- Suh, S., & Trabasso, T. (1993). Inferences during reading: Converging evidence from discourse analysis, Talk-aloud protocols, and recognition priming. *Journal of memory and language*, 32, 279-300.
- Van de Wiel, M.W.J. (1997). Knowledge encapsulation: Studies on the development of medical expertise. Doctoral dissertation. Wageningen: Ponsen & Looijen.
- Van de Wiel, M.W.J., Boshuizen, H.P.A., Schmidt, H.G (1998). A failure to reproduce the intermediate effect in clinical case recall. *Academic Medicine*, 73, 894-900.

Samenvatting & Conclusies

In dit proefschrift staan de verschillen centraal in de aard van de kennis die specialisten in de geneeskunde gebruiken wanneer ze problemen binnen en buiten hun vakgebied moeten oplossen.

De ontwikkeling van expertise heeft in de cognitieve psychologie de laatste decennia veel aandacht gekregen. Zo is er veel onderzoek verricht naar de ontwikkeling van expertise in de geneeskunde. Dit onderzoek was met name gericht op het achterhalen van de kennisstructuur, zoals die bij personen van verschillend expertiseniveau aanwezig is. Meestal werden beginnende en gevorderde studenten geneeskunde met medisch experts vergeleken in de wijze waarop ze een klinische casus diagnostiseerden. Deze casus bevatte informatie over de klachten van de patiënt, de anamnese, laboratoriumdata, bevindingen van het lichamelijk onderzoek en de uitkomsten van aanvullend onderzoek (bijv. ECG's en röntgenfoto's). De proefpersonen werd gevraagd een casus te bestuderen om achtereenvolgens een diagnose te stellen, op te schrijven wat ze zich van de casus konden herinneren (de zogenaamde "free recall") en tenslotte een pathofysiologische verklaring te geven voor de verschijnselen zoals beschreven in de casus. Wat deze studies met enige regelmaat hebben laten zien, is het zogenaamde "intermediate effect". Dat wil zeggen, gevorderde studenten ("intermediates") herinnerden zich meer van de beschrijving van een casus dan beginnende studenten en medisch experts (bijv. huisartsen of internisten). Bovendien gaven gevorderde studenten veel uitgebreidere pathofysiologische verklaringen voor de verschijnselen die in een casus beschreven werden dan de andere groepen. Deze tegenintuïtieve bevindingen zijn moeilijk te verklaren op basis van de meeste theorieën die de ontwikkeling van expertise beschrijven, omdat deze theorieën meestal een lineair verband tussen de prestaties met de mate van expertise voorspellen: experts zouden dus de beste prestaties moeten leveren. Teneinde deze verschillen te verklaren, hebben Schmidt en Boshuizen (1992, 1993) de *encapsulatie-theorie* ontwikkeld. Deze theorie stelt dat er kwalitatieve veranderingen hebben plaatsgevonden tijdens het traject van beginnend student geneeskunde tot ervaren medisch specialist. Studenten geneeskunde vergaren tijdens de eerste jaren van hun opleiding met name basale biomedische kennis en worden amper geconfronteerd met de klinische manifestaties van aandoeningen. Zij kunnen derhalve nog niet beschikken over uitgebreide kennis betreffende de wijze waarop een ziekte zich manifesteert. Medisch experts daarentegen, kennen ook de klinische zijde van een aandoening en zullen deze kennis ook toepassen bij de evaluatie van een patiënt. Dat wil niet zeggen dat experts hun biomedische kennis vergeten zijn. Schmidt en Boshuizen beweren veeleer, dat hun biomedische kennis geïntegreerd is geraakt in hun klinische kennis. Dit integratieproces hebben deze onderzoekers *kenniscapsulatie* genoemd. Door middel van de theorie van kenniscapsulatie wordt het

mogelijk het 'intermediate effect' te verklaren. Experts introduceren veel encapsulerende concepten in hun "recall" en pathofysiologische verklaringen, terwijl gevorderde studenten veel gedetailleerder te werk gaan, zonder dat er sprake is van encapsulaties. Wanneer de data van experts en gevorderde studenten puur kwantitatief benaderd worden, dan worden dus betere prestaties bij gevorderde studenten aangetroffen: zij produceren immers meer gegevens over de getoonde casus. Kijkt men evenwel naar de kwaliteit van de verklaringen, dan wordt duidelijk dat experts door middel van encapsulaties veel beter in staat zijn het kaf van het koren te scheiden.

De geëncapsuleerde kennisstructuur maakt het mogelijk dat experts snel en accuraat problemen binnen hun vakgebied oplossen. Het is evenwel niet duidelijk wat er zal gebeuren wanneer experts geconfronteerd worden met problemen die buiten hun domein van expertise liggen en die zelden of nooit worden waargenomen. Kunnen experts, of beter gezegd *subexperts* (d.w.z. medisch specialisten die problemen buiten hun vakgebied moeten oplossen), ook onder zulke omstandigheden een beroep doen op hun geëncapsuleerde kennis? Dit brengt ons tot de vraag die in deze studie centraal staat: *wat is de aard van de kennis die experts gebruiken als ze non-routine problemen moeten diagnostiseren?*

Uitgaande van de theorie van kennisencapsulatie kunnen we een aantal voorspellingen doen met betrekking tot de verschillen, zoals die verwacht worden op te treden tussen experts die problemen binnen of buiten hun vakgebied moeten oplossen. Experts evalueren problemen buiten hun vakgebied minder snel en accuraat, en hun "recall" en pathofysiologische verklaringen zullen uitgebreider zijn, omdat ze buiten hun domein van expertise in onvoldoende mate over geëncapsuleerde concepten kunnen beschikken.

Vier studies zijn uitgevoerd teneinde de rol van geëncapsuleerde kennis binnen en buiten het domein van de medisch specialist te onderzoeken. In hoofdstuk 2 wordt een studie beschreven waarin cardiologen en pulmonologen vier casus moesten bestuderen. Twee casus waren binnen het vakgebied van de cardiologie en twee binnen de pulmonologie. Nadat de specialisten de casus gelezen hadden, dienden ze een diagnose te stellen, op te schrijven wat ze zich nog van de casus konden herinneren ("free recall") en tenslotte een pathofysiologische verklaring te geven, waarin ze de verschijnselen zoals beschreven in de casus moesten verklaren. De resultaten lieten zien dat experts binnen hun vakgebied sneller waren in het bestuderen van de casus en een hogere diagnostische accuratesse behaalden, verschillen in "recall" en pathofysiologie waren evenwel afwezig. Deze laatste bevinding wijkt af van de voorspelling zoals die gemaakt is op basis van de theorie van kennisencapsulatie. Een mogelijke verklaring voor deze bevinding werd gevonden in de hoge complexiteit van de casuïstiek. Experts waren niet in staat de casus op een "routinematige" manier te verwerken (d.w.z. op een geëncapsuleerde wijze) en hebben derhalve de casus verwerkt door terug te vallen op hun biomedische kennis - overeenkomstig de wijze waarop subexperts de casus verwerkt hebben. Een alternatieve verklaring die voor de bevindingen gegeven werd, is

afkomstig van Patel en Groen (Patel & Groen, 1991a, 1991b). Volgens deze onderzoekers hebben experts en subexperts een vorm van *generieke expertise* ontwikkeld. Dat wil zeggen, experts en subexperts hebben schema's verworven die tot een bepaalde aandoening behoren en stellen hen in staat de essentiële informatie van een casus te onthouden. In tegenstelling tot subexperts hebben domeinexperts ook *specifieke expertise* ontwikkeld, welke hen in staat stelt de onopvallende en onverklaarbare symptomen in verband te brengen met de diagnose, hetgeen uiteindelijk resulteert in een hogere diagnostische accuratesse. Een derde mogelijkheid die besproken werd, had betrekking op de relatief geringe "afstand" tussen de specialismen van cardiologie en pulmonologie. Dit is met name in Nederland het geval waar cardiologen en pulmonologen een deel van hun opleiding gezamenlijk doorlopen en derhalve regelmatig geconfronteerd worden met cardiologische en pulmonologische problemen. Het is niet uit te sluiten dat kennisencapsulatie voor beide specialismen reeds in een vrij vroeg stadium optreedt tijdens hun assistentschap.

De studie beschreven in Hoofdstuk 3 heeft geprobeerd de problemen van de complexiteit van de casuïstiek en de geringe afstand tussen de specialismen zoals beschreven in Hoofdstuk 2 op te lossen. Twee casusbeschrijvingen zijn gebruikt welke reeds in voorgaande studies gebruikt zijn (onder andere door Patel & Groen, 1986; Schmidt & Boshuizen, 1993; Van de Wiel, Schmidt, & Boshuizen, 1998). Wederom waren de proefpersonen medisch specialisten, die gevraagd werden twee casus te bestuderen, op te schrijven wat ze zich van de casus herinnerden en tenslotte een pathofysiologische verklaring te geven. In tegenstelling tot het eerste experiment, was er maar een groep specialisten die de rol van expert vervulde (cardiologen). De andere groep van specialisten (neurologen) vervulde de rol van subexperts. Beide groepen hadden drie minuten voor het bestuderen van iedere casus. De resultaten van deze studie waren in overeenstemming met de voorspellingen zoals die gemaakt zijn op basis van de encapsulatietheorie. Experts waren accurater in hun diagnose dan subexperts en bovendien gebruikten experts meer geëncapsuleerde concepten in hun "recall" en pathofysiologische verklaringen. Hoewel deze bevinden goed verklaard kunnen worden op basis van de theorie van kennisencapsulatie, lijken enige kanttekeningen gerechtvaardigd. Ten eerste, de diagnostische accuratesse van de experts was niet erg hoog (ongeveer 73 % van de maximumscore) en maar twee cardiologen gaven een volledig correcte diagnose voor beide casus. Dit is een aanwijzing dat de casuïstiek bestond uit complexe problemen voor de cardiologen en het is derhalve niet uit te sluiten dat deze experts niet in staat waren deze casus volledig routinematig te verwerken. Ten tweede, het aantal geëncapsuleerde concepten in de "recall" van de experts was laag (17 % van het totaal aantal herinnerde proposities). Weliswaar gebruikten experts significant meer geëncapsuleerde concepten dan subexperts, toch ondermijnt deze bevinding de bijdrage van geëncapsuleerde kennis aan de verschillen zoals die gevonden zijn in het totaal aantal gereproduceerde proposities tussen experts en subexperts. Dat wil zeggen, als men de proportie geëncapsuleerde concepten nader bekijkt, dan kan men

niet zonder meer concluderen dat de “recall” van experts gedomineerd wordt door geëncapsuleerde concepten. Tenslotte, het aantal gebruikte *concepten* en *links* door experts bij hun pathofysiologische verklaringen wijkt niet significant af van die van subexperts. Daar staat tegenover dat experts grofweg twee keer zoveel geëncapsuleerde concepten in hun verklaringen gebruikten dan subexperts. Dit is wederom een aanwijzing dat de uitgebreidheid van de pathofysiologische verklaringen relatief onafhankelijk is van het aantal geëncapsuleerde concepten. Dit gegeven is niet onbelangrijk, omdat het zogenaamde “intermediate effect” bij klinische casusstudies verklaard werd door met name de invloed van geëncapsuleerde concepten.

De studie beschreven in Hoofdstuk 4 onderzocht de relatie tussen geëncapsuleerde kennis en het “intermediate effect”. De proefpersonen waren tweede- en zesdejaars studenten geneeskunde en neurologen en ze moesten vier casusbeschrijvingen bestuderen, die buiten het vakgebied van de neurologen lagen. Vervolgens moesten ze een diagnose stellen, opschrijven wat ze zich van de casus herinnerden, en tenslotte een pathofysiologische verklaring geven. De resultaten laten een “intermediate effect” zien voor “recall” en pathofysiologische verklaringen. Deze resultaten suggereren dat experts relatief ongevoelig zijn voor de aard van de gebruikte casuïstiek. Diagnostische accuratesse daarentegen, liet een lineair verband zien met de mate van expertise. Deze studie werpt een nieuw licht op het “intermediate effect,” omdat dit fenomeen zowel binnen als buiten het expertisedomein van een expert kan optreden. Daarenboven plaatst deze studie en die beschreven in Hoofdstuk 2 en 3, vraagtekens bij de rol van geëncapsuleerde kennis als belangrijkste constituent van het “intermediate effect”. De vraag die nu beantwoord dient te worden is, wat zijn de oorzaken van het intermediate effect als geëncapsuleerde kennis niet de enige factor van betekenis is? Een mogelijke verklaring kan gevonden worden in de verschillen in motivatie of doelstellingen tussen experts en studenten. Schmidt (1983, 1984) bijvoorbeeld, heeft een *cognitive set-point* hypothese voor-gesteld. Hiermee bedoeld hij dat tijdens het verwerkingsproces van tekstuele informatie een interne standaard (“set-point”) wordt bepaald, welke aangeeft wanneer er voldoende tijd is besteed aan het verwerken van informatie met het doel het in een latere fase te reproduceren. Er wordt verondersteld dat de set-point bepaald wordt door de inschatting van de lezer over de minimale hoeveelheid begrip, zijn of haar doelstellingen, motivatie en externe druk (bijv. de consequenties die verbonden zijn met het niet adequaat kunnen reproduceren van de informatie). Wanneer het set-point niveau voor een bepaalde informatie-eenheid bereikt is, dan zal de lezer doorgaan met het verwerken van nieuwe informatie totdat de set-point voor alle eenheden bereikt is of de tijd op is. Wanneer proefpersonen aan een onderzoek in hoge mate beschikken over voorkennis - zoals de experts in het huidige onderzoek - dan hebben ze minder tijd nodig hun set-point te bereiken voor de informatie (Machiels-Bongaerts, 1993). Teneinde de plausibiliteit van geëncapsuleerde kennis te doorgronden en de problemen die verbonden zijn met de “traditionele” aanpak van casus studies te

voorkomen, is er voor een meer directe benadering gekozen om de casusrepresentatie van experts en subexperts te onderzoeken. Hoofdstuk 5 beschrijft twee zogenaamde *primingexperimenten* welke wezenlijk afwijken van de traditionele benadering. In plaats van het opschrijven van hetgeen men zich herinnert van de casus of een verklaring te geven voor de verschijnselen, werd de proefpersonen gevraagd zo snel en accuraat mogelijk te beslissen of een medisch concept dat gepresenteerd werd op een computerscherm gerelateerd is aan de casus die ze net gelezen hebben. De casus heeft nu de functie van een *prime* voor alle kennis die geactiveerd wordt tijdens het diagnostische proces. Deze pre-activatie van medische concepten door de casus leidt vervolgens tot facilitatie van het evaluatieproces dat gebruik maakt van deze informatie in het geheugen (het zogenaamde *semantic priming effect*, Collins & Loftus, 1975). Wanneer encapsulatietheorie gelijk heeft in de aanname dat medisch experts een representatie van een casus binnen hun expertisedomein opbouwen op basis van geëncapsuleerde concepten, dan mag men verwachten dat experts ook sneller en accurater reageren dan studenten of subexperts, als ze moeten beslissen of een geëncapsuleerd concepten gerelateerd is met de casus. Van studenten en subexperts wordt immers verwacht dat ze een representatie van een casus opbouwen die voornamelijk bestaat uit biomedische concepten (Schmidt & Boshuizen, 1992, 1993). Het eerste experiment bestond uit twee casusbeschrijvingen, die ieder binnen een minuut bestudeerd moesten worden. De helft van de proefpersonen kreeg de correcte diagnose vooraf, de andere helft moest zelf een diagnose stellen. Vervolgens moesten alle proefpersonen drie typen van items op hun gerelateerdheid tot de casus beoordelen: letterlijk beschreven symptomen, afgeleide geëncapsuleerde items (waren dus geen onderdeel van de casusbeschrijving), en filler items (waren totaal irrelevant voor de casus). De resultaten waren overeenkomstig de encapsulatietheorie: experts waren sneller en accurater dan studenten en subexperts in het evalueren van de items. Bovendien waren experts ook sneller en accurater in het evalueren van de letterlijk beschreven symptomen. Een tegenovergestelde relatie was het geval voor de subexperts. Zij waren sneller in het beoordelen van de symptoom-items dan de geëncapsuleerde items. Studenten, tenslotte, lieten geen "voorkeur" zien voor een bepaald type item. Het was opmerkelijk dat experts die een diagnose moesten stellen een relatief lage accuratesse behaald hebben (ongeveer 50 % van de maximum score) - niettemin significant hoger dan beide andere groepen. Dit kan een aanwijzing zijn dat experts, die vooraf geen correcte diagnose kregen, de casus niet geheel in geëncapsuleerde modus verwerkt hebben. Immers, verschillen tussen experts die wel of niet op de hoogte waren gesteld van de correcte diagnose zouden dan niet waarschijnlijk zijn. De invloed van het geven van een correcte diagnose heeft overigens een gelijke invloed op alle groepen. Studenten, subexperts, en experts waren sneller op beide typen van items wanneer ze geïnformeerd waren over de correcte diagnose dan wanneer dat niet het geval was. Het geven van de correcte diagnose heeft een vergelijkbare invloed op de geëncapsuleerde items als op de symptomen.

Hoewel het eerste primingexperiment een beter inzicht heeft opgeleverd in de rol van geëncapsuleerde kennis en die van symptomen in de casusrepresentatie van de proefpersonen, blijft de rol van biomedische kennis vooralsnog onduidelijk. Het tweede experiment van Hoofdstuk 5 heeft derhalve de rol van biomedische kennis en geëncapsuleerde kennis in de casusrepresentatie onderzocht. Schmidt en Boshuizen (1992, 1993) veronderstellen dat medisch experts terugvallen op hun biomedische kennis als ze geconfronteerd worden met problemen die buiten hun vakgebied liggen. Vierdejaars studenten geneeskunde, subexperts, en experts werd gevraagd een tekst te bestuderen welke de gangbare symptomen beschrijft die verbonden zijn met *hartfalen* (Een cardiologische aandoening die ook wel *decompensatio cordis* wordt genoemd). Volgens Schmidt en Boshuizen activeren de symptomen beschreven in de casus specifieke geëncapsuleerde concepten in het langetermijngeheugen van een expert. Studenten en subexperts daarentegen, kunnen geen gebruik maken van geëncapsuleerde kennis en moeten derhalve terugvallen op hun biomedische kennis teneinde de verschijnselen in de casus te begrijpen. Wederom werden drie typen items gebruikt: geëncapsuleerde items, biomedische items en filler items. Zowel de geëncapsuleerde items als de biomedische items werden niet in de tekst vermeld. De resultaten lieten zien dat experts sneller waren dan beide andere groepen in het evalueren van geëncapsuleerde items. Studenten en subexperts lieten dezelfde tendens zien in hun responsietijden als de experts, hoewel ze niet zo snel waren als de experts. Dat wil zeggen, studenten en subexperts evalueerden de geëncapsuleerde items sneller dan de biomedische items. Er waren overigens geen verschillen in het gemiddeld aantal fouten tussen de verschillende expertiseniveaus. De bevindingen van beide primingexperimenten geven ons een nieuw perspectief op de beschreven studies die uitgaan van het traditionele onderzoeksparadigma. Ten eerste suggereren deze primingstudies dat zowel experts als subexperts en zelfs studenten beschikken over geëncapsuleerde kennis en dus een casus in feite op dezelfde wijze verwerken. Geëncapsuleerde kennis is dus geen uniek kenmerk van de kennisstructuur van een expert. Daar staat tegenover dat de responsietijden van de experts duidelijk laten zien, dat geëncapsuleerde concepten bij met name experts een speciale plaats innemen in hun medische kennis. Ten tweede, het expertise-onderzoek door middel van priming lijkt een zeer geschikte benadering om verschillen te achterhalen zoals die bestaan in de casusrepresentaties van experts en subexperts.

Conclusies

De eerste conclusie die getrokken kan worden uit de bevindingen is dat ze de theorie van kennisencapsulatie grotendeels ondersteunen. Met name de "primingstudies" ondersteunen de unieke positie van geëncapsuleerde concepten in de casusrepresentatie van een medisch expert. Experts zijn sneller en meer accuraat in het evalueren van geëncapsuleerde concepten dan subexperts en studenten genees-

kunde. Geëncapsuleerde concepten zijn evenwel geen unieke karakteristiek van de expert. Subexperts en studenten beschikken ook over geëncapsuleerde concepten en gebruiken deze, zij het in mindere mate dan de experts, in hun casusrepresentatie. Het wezenlijke verschil tussen experts, subexperts en studenten is derhalve niet het al dan niet kunnen beschikken over geëncapsuleerde kennis, maar de mate waarin geëncapsuleerde concepten geactiveerd worden tijdens het verwerken van casus-informatie. De bevindingen suggereren dat de geëncapsuleerde concepten bij experts in hogere mate geactiveerd worden dan dat het geval is bij studenten of subexperts.

Een tweede conclusie die getrokken kan worden uit de beschreven studies is dat maten zoals verwerkingssnelheid van een casus en diagnostische accuratesse stabiele en betrouwbare maten zijn voor het detecteren van verschillen tussen de verschillende expertiseniveaus. Een bevinding die gecorrobooreerd wordt door voorgaande casusstudies (e.g., Claessen & Boshuizen, 1985; Schmidt & Boshuizen, 1993; Van de Wiel, 1997). “Free recall” en pathofysiologische verklaringen daarentegen, zijn meer problematische maten. De studie beschreven in Hoofdstuk 2 bijvoorbeeld, laat geen significante verschillen zien tussen experts en subexperts, hoewel deze wel voorspeld waren. Hoofdstuk 3 daarentegen beschrijft een studie met experts en subexperts die vrijwel volledig overeenkwam met de voorspellingen afgeleid van de encapsulatietheorie. Wanneer men nu de verschillende uitkomsten in ogenschouw neemt, kan men concluderen dat “recall” en pathofysiologische verklaringen niet geheel betrouwbare maten zijn om de verschillen tussen experts en subexperts te onderzoeken. Primingstudies zoals beschreven in Hoofdstuk 5 lijken in dit opzicht meer hoopvol.

Een derde conclusie die getrokken kan worden is dat het terugvallen van experts buiten hun specialisme niet louter en alleen toegeschreven kan worden aan het al dan niet activeren van geëncapsuleerde kennis. Op de eerste plaats, omdat de studie beschreven in Hoofdstuk 2 geen verschillen heeft laten zien tussen experts en subexperts in “recall” en pathofysiologie. Op de tweede plaats, omdat de studie beschreven in Hoofdstuk 4 heeft laten zien dat een “intermediate effect” ook buiten het specialisme van de expert kan optreden. Alternatieve verklaringen zoals die bijvoorbeeld gegeven worden door Patel en Groen (1991a, 1991b) lijken derhalve zeer plausibel.

Dankwoord

Wie had het kunnen denken dat een straatschoffie van de Maastrichtse wijk Belfort het ooit zover zou schoppen? De leraren van mijn lagere school in ieder geval niet: die deelden ieder jaar mijn rapport hoofdschuddend uit.

Gelukkig heb ik de lange weg niet alleen hoeven af te leggen. In het bijzonder ben ik oneindig veel dank verschuldigd aan wijlen Nico Hougardy: mijn (schaak)leraar. Hij was het die veel meer in mij zag dan mijn toenmalige leraren van de lagere school voor mogelijk hielden. “Meneer Hougardy,” U heeft gelijk gekregen.

Na het voltooien van de opleiding cognitiewetenschappen in Nijmegen werd ik Onderzoeker in Opleiding (OiO) aan de Universiteit Maastricht. Ik wil Els Boshuizen danken, omdat zij mij tijdens de eerste jaren van mijn onderzoek de fijne kneepjes van het vak heeft bijgebracht. Zij was het die mijn vaak nog wilde ideeën in goede banen wist te leiden.

De laatste twee jaren van mijn OiO-schap heb ik doorgebracht bij de Faculteit der Psychologie en nam Henk Schmidt me persoonlijk onder zijn hoede. Tijdens deze periode, die in feite nog tot de dag van vandaag voortduurt, hebben we samen zoveel nieuwe onderzoeksideeën gegenereert, dat het expertise-onderzoek voor de komende jaren is veiliggesteld. Ik ben hem zeer dankbaar voor zijn rotsvast vertrouwen dat hij in mij had: het gaf mij de moed iedere (nieuwe) uitdaging aan te gaan.

Daarnaast wil ik mijn collega's van vakgroepen O&O (Onderwijsontwikkeling & Onderwijsresearch) en psychologie bedanken voor het creëren van een goede en vruchtbare werksfeer. In het bijzonder wil ik Fred Paas danken voor zijn vaak pittige maar nooit ongegronde kritiek op mijn pennevruchten.

Het is vaak ter ziele gedragen om vervolgens weer als een feniks uit zijn as te herrijzen: het leesgroepje bij O&O. Eugène Custers, Herco Fontein, Rina Vaatstra, Margje van de Wiel, ik dank jullie allemaal voor de vaak produktieve discussies. Bovendien ben ik Margje dank verschuldigd voor de casuïstiek die zij mij vaak belangeloos beschikbaar stelde.

Het ontwikkelen van medische casuïstiek is een wezenlijk onderdeel van het expertise-onderzoek. Gelukkig heb ik een aantal medisch specialisten in het azM bereid gevonden mij bij deze belangrijke klus te helpen. In het bijzonder ben ik Geertjan Wesseling, Gerard Linssen, en Nicolaas Schaper hiervoor dankbaar en natuurlijk ook al die specialisten, die ondanks hun overvolle agenda's, toch de tijd vonden aan dit onderzoek deel te nemen.

Mijn werkzaamheden als OiO werden aanzienlijk vergemakkelijkt door een aantal student-assistenten. In chronologische volgorde hebben de volgende studenten mij terzijde gestaan: Ali Mert, Astrid Wanders, Bart Hugén en Katrien Pouls.

Tenslotte wil ik mijn familie en vrienden danken die ieder op hun eigen wijze aan de verwezenlijking van dit proefschrift hebben bijgedragen. Op de eerste plaats mijn ouders, die zich nog steeds zorgen maken of het wel allemaal goed komt. Jolanda en René voor de “cultuurtochten” en hun sponsoring. In het bijzonder Marianne van den Hurk, die na twee jaar op mijn kamer te hebben gezeten, nog steeds mijn beste vriendin is. Tenslotte, Wilco Tewinkel, Pascal van Gerven en Jettie Hoonhout die stevast onderdeel uitmaken van de kroegploeg. Mijn dank voor de noodzakelijke ontspanning.

Remy Rikers

Maastricht, November 1997

Curriculum Vitae

Remy Rikers was born on August 26, 1967 in Maastricht. He attended secondary education at the Hendrik van Veldeke College in Maastricht. From 1989 till 1990 he worked at the library for the blind (NLBB) in The Hague as a substitute for military service. He studied cognitive science and philosophy at the University of Nijmegen (KUN) from 1990 till 1994. In 1995 he started his Ph.D. research on the development of medical expertise. The first two years he spent at the department of Educational Development and Research (O&O), the second part of his research project he worked at the Faculty of Psychology. Since March 1999 he works as a lecturer at the department of experimental psychology.